

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2002, 21:01:30 ; Search time 562.42 Seconds
(without alignments)
11817.104 Million cell updates/sec

Title: US-09-729-674-1

Perfect score: 3871

Sequence: 1 ttctctctccctccctt.....ataaaaaaaaaaaaaa 3871

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq-032802:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3870.2	100.0	3871	20	AAH60801 Human secreted pro
2	3870.2	100.0	3871	22	AAH59207 Human CDNA encodin
3	3870.2	100.0	3871	24	AAH90876 Human polynucleoti
4	2051.2	53.0	2549	21	AAH16225 Human prostate can
5	1955.4	50.5	1961	22	AAH14525 Human CDNA sequenc
6	1474.4	38.1	1586	21	AAH31138 Human colon cancer
7	695.6	18.0	818	22	AAH06808 Human CDNA clone (
8	466.4	12.0	618	21	AAZ80131 Human colon cancer
9	406.4	10.5	408	21	AAH30585 Human colon cancer

SEA ID NO: 1
Database: N.Geneseq-032802
AC NO: AAH60801

10	371.6	9.6	401	22	AAH50786 Human tumour assoc
11	369.4	9.5	400	16	AAH21281 Human gene signatu
12	355.6	9.2	407	22	AAH50808 Human tumour assoc
13	340	8.8	351	22	AAH11840 Human CDNA clone (
14	337.2	8.7	373	21	AAH43677 Human secreted exp
15	258.6	6.7	279	23	AAH58007 CDNA #683 encoding
16	239.2	6.2	6670	22	AAH02846 Human reproductive
17	239.2	6.2	7713	22	AAH02845 Human reproductive
18	238.4	6.2	8658	22	AAH69231 Human immune/haema
19	238	6.1	295	22	AAH82995 Human immune/haema
20	235.4	6.1	32351	21	AAH21307 Human low adenosin
21	235.4	6.1	32351	21	AAH35185 Human adenosine re
22	235.4	6.1	40298	21	AAH21311 Human low adenosin
23	235.4	6.1	40298	21	AAH35189 Human adenosine re
24	234.6	6.1	4275	22	AAH19408 Human nervous syst
25	234.6	6.1	4883	22	AAH79410 Human immune/haema
26	234.6	6.1	16997	22	AAH80033 Human immune/haema
27	234.6	6.1	16997	22	AAH86146 Human reproductive
28	233.8	6.0	9620	22	AAH06207 Human liver cell s
29	233.4	6.0	6138	22	AAH57456 Human CDNA clone (
30	233.2	6.0	763	22	AAH06411 Human CDNA sequenc
31	232.6	6.0	2263	22	AAH14438 Human immune/haema
32	232.6	6.0	5979	22	AAH78985 Human immune/haema
33	232.6	6.0	5979	22	AAH80951 Human cervical can
34	232.2	6.0	532	22	AAH71264 Human cervical can
35	232.2	6.0	567	22	AAH69839 Human polynucleoti
36	232	6.0	3775	22	AAH61061 Human immune/haema
37	231.4	6.0	1449	22	AAH86057 Human immune/haema
38	231.4	6.0	1450	22	AAH86062 Human nervous syst
39	230.8	6.0	2847	22	AAH18224 Human immune/haema
40	230.6	6.0	2108	22	AAH82073 Human immune/haema
41	230.6	6.0	2108	22	AAH76677 Human immune/haema
42	230.6	6.0	2108	22	AAH76678 Human immune/haema
43	230.2	5.9	146	22	AAH77915 Human immune/haema
44	230.2	5.9	1146	22	AAH90464 Human digestive sy
45	230	5.9	13559	22	AAH15144 Human nervous syst

ALIGNMENTS

RESULT 1	
AAH60801	
ID	AAH60801 standard; DNA: 3871 BP.
XX	
AC	AAH60801:
XX	
DT	09-AUG-1999 (first entry)
XX	
DE	Human secreted protein encoding DNA (clone bd306-7).
XX	
KW	Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW	nutritional activity; cytokine; cell proliferation; immune stimulation;
KW	hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW	anti-inflammatory; tumour invasion; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO926961-A1.
XX	
PD	03-JUN-1999.
XX	
PF	24-NOV-1998; 98WO-US25149.
XX	
PR	23-NOV-1998; 98US-0197886.
XX	
PR	26-NOV-1997; 97US-0066804.
XX	
PA	(GEMV) GENETICS INST INC.
XX	
PI	Agostino M, Clark HF, Collins-Racie LA, Evans C;
XX	Feichel K, Jacobs K, Lavallie ER, McCoy DM, Werberg D;
XX	Steininger RJ, Treacy M, Wong GG;

|||||
Db 1681 ccaggcgtgtcgaactcttgacttgactcaatgacatccactgtgccttggcctccccaagtgc 1740
QY 1741 tgggattacaggcgctgagccacacatgcgcggcctcttccacacttaacacttctct 1800
Db 1741 tgggattacaggcgctgagccacacatgcgcggcctcttccacacttaacacttctct 1800
QY 1801 tatctacacatctgttctcaacactcatccctgtcttctctcaatgttcaacactgtcttc 1860
Db 1801 tatctacacatctgttctcaacactcatccctgtcttctctcaatgttcaacactgtcttc 1860
QY 1861 ccaatgttcaatgctgccttcttcaacatttggtttgaaggcgagctctctgtgcttg 1920
Db 1861 ccaatgttcaatgctgccttcttcaacatttggtttgaaggcgagctctctgtgcttg 1920
QY 1921 ttttttggttttcccgagaaatcagratatttttaataagaanaaatcttccatgaa 1960
Db 1921 ttttttggttttcccgagaaatcagratatttttaataagaanaaatcttccatgaa 1960
QY 1981 gatgataattgtgaaacccctctgtgcttattgtcttccagattttagctccttc 2040
Db 1981 gatgataattgtgaaacccctctgtgcttattgtcttccagattttagctccttc 2040
QY 2041 tccccatccgggaaagaatggctggaagacataggtctaaattctccagcctcaacatgtgc 2100
Db 2041 tccccatccgggaaagaatggctggaagacataggtctaaattctccagcctcaacatgtgc 2100
QY 2101 ttaactgtgtctgactgttccacttcttaagcaccacttgaaacaaagtttgatagagag 2160
Db 2101 ttaactgtgtctgactgttccacttcttaagcaccacttgaaacaaagtttgatagagag 2160
QY 2161 tgtagagctgcagaagaatgttgcttcttgcccaacttgcacatcccaaatcaaaagtgtgag 2220
Db 2161 tgtagagctgcagaagaatgttgcttcttgcccaacttgcacatcccaaatcaaaagtgtgag 2220
QY 2221 ccgactcccaattgaggaacaaatgtcttgaatlaagtlccaggttgcgaaaggaaagaaag 2280
Db 2221 ccgactcccaattgaggaacaaatgtcttgaatlaagtlccaggttgcgaaaggaaagaaag 2280
QY 2281 caagagctgtctagtatttcatcttcttccagtaataataattatgagtactactgtgtgc 2340
Db 2281 caagagctgtctagtatttcatcttcttccagtaataataattatgagtactactgtgtgc 2340
QY 2341 taagcatctgcacccgggagactagagatatcttcacagaaatacaggaagaagttccctgtgc 2400
Db 2341 taagcatctgcacccgggagactagagatatcttcacagaaatacaggaagaagttccctgtgc 2400
QY 2401 catggaagcttacaattctacaggggagaaagatagccaatacataggaataataataatc 2460
Db 2401 catggaagcttacaattctacaggggagaaagatagccaatacataggaataataataatc 2460
QY 2461 aaggtatcatgtatgtatgaatttgcctgttgcgagaaataaagcagggagggagtgaagaa 2520
Db 2461 aaggtatcatgtatgtatgaatttgcctgttgcgagaaataaagcagggagggagtgaagaa 2520
QY 2521 tctctgagatagagcctcgaattttaatgggctcctcaactggaatggaacttggagaga 2580
Db 2521 tctctgagatagagcctcgaattttaatgggctcctcaactggaatggaacttggagaga 2580
QY 2581 gacgtttagggaagtgtgacttcttgacaagcatctccagcgagaggaacaaagtgtgcactg 2640
Db 2581 gacgtttagggaagtgtgacttcttgacaagcatctccagcgagaggaacaaagtgtgcactg 2640
QY 2641 ccccaagaatgagaacttgcctctacgtgttcaagaaagagagagagccaaagcagagtcg 2700
Db 2641 ccccaagaatgagaacttgcctctacgtgttcaagaaagagagagagccaaagcagagtcg 2700
QY 2701 tgggcaagggctagaatggaagagagcggtctgggagagacaggtgtgtgaaggccttgg 2760
Db 2701 tgggcaagggctagaatggaagagagcggtctgggagagacaggtgtgtgaaggccttgg 2760
QY 2761 ctctctcaagtgaatggaacccactggaagggttctgaacagagaagtgtccttgatgat 2820
|||||

Db 2761 ctctctcaagtgaatggaacccactggaagggttctgaacagagaatgtccttgatgat 2820
QY 2821 ttatatttgcagaaggtcatctctagctgcacatattgtgaaataactttagtgcagaggc 2880
Db 2821 ttatatttgcagaaggtcatctctagctgcacatattgtgaaataactttagtgcagaggc 2880
QY 2881 agaaagagagagagacactgttaggaagctactgcagagttccagagcttggcctgggc 2940
Db 2881 agaaagagagagagagacactgttaggaagctactgcagagttccagagcttggcctgggc 2940
QY 2941 cacaagcaacagcagtggtccaataatctagatttatttgaagaagccaatagatctgc 3000
Db 2941 cacaagcaacagcagtggtccaataatctagatttatttgaagaagccaatagatctgc 3000
QY 3001 tgaagtttaagtgtgagtgctgaagagaaggaagtttaagtatgacatlaagtttttg 3060
Db 3001 tgaagtttaagtgtgagtgctgaagagaaggaagtttaagtatgacatlaagtttttg 3060
QY 3061 gctctgaatagcagaagaagatggaatlaaccagttaactgaataataggaagatggctgggt 3120
Db 3061 gctctgaatagcagaagaagatggaatlaaccagttaactgaataataggaagatggctgggt 3120
QY 3121 aagtaaggaattgtgtcacaagcagctgtctgtgttgcgaaatggagaggtctgtgcga 3180
Db 3121 aagtaaggaattgtgtcacaagcagctgtctgtgttgcgaaatggagaggtctgtgcga 3180
QY 3181 aatcaagtgagaatcttctcaggttcaggtctgaagcagaactcgaagcagaaggtctgaa 3240
Db 3181 aatcaagtgagaatcttctcaggttcaggtctgaagcagaactcgaagcagaaggtctgaa 3240
QY 3241 tgcacttggttatctgttgggggtcctcagaagagaaacctgtgaagccttatacagctc 3300
Db 3241 tgcacttggttatctgttgggggtcctcagaagagaaacctgtgaagccttatacagctc 3300
QY 3301 atttattgtgctgtgagaagttctctggagatgtgtgtatcatcttgaagcaagttactca 3360
Db 3301 atttattgtgctgtgagaagttctctggagatgtgtgtatcatcttgaagcaagttactca 3360
QY 3361 gttgagggcagaatctctgggaaagaagcctgaagcactcggacatccacatgctgttag 3420
Db 3361 gttgagggcagaatctctgggaaagaagcctgaagcactcggacatccacatgctgttag 3420
QY 3421 tgtgttgggggttgggggtcctcggcactgtgtgtgaagagatctgcagggaccca 3480
Db 3421 tgtgttgggggttgggggtcctcggcactgtgtgtgaagagatctgcagggaccca 3480
QY 3481 ggcgcacctactgaaacatcagcatgtctcagttgacatttaagccaatgcaagctggaagg 3540
Db 3481 ggcgcacctactgaaacatcagcatgtctcagttgacatttaagccaatgcaagctggaagg 3540
QY 3541 cacttgaatgtctctgtatgattactctgaagaagcaacgaagaagccatgtgagctcc 3600
Db 3541 cacttgaatgtctctgtatgattactctgaagaagcaacgaagaagccatgtgagctcc 3600
QY 3601 ttgggctcctctgggaaatgggaaatcgcgaagagactggaaggaattactctaagctc 3660
Db 3601 ttgggctcctctgggaaatgggaaatcgcgaagagactggaaggaattactctaagctc 3660
QY 3661 agagaagaacagagagtggtgtctctggaagcttgagcttcttataccaactcatc 3720
Db 3661 agagaagaacagagagtggtgtctctggaagcttgagcttcttataccaactcatc 3720
QY 3721 ccttctccaataagccaactgtgtagttggccctccgggtgttgaaggcaagaagaga 3780
Db 3721 ccttctccaataagccaactgtgtagttggccctccgggtgttgaaggcaagaagaga 3780
QY 3781 aaggaagcgttctgggaaacaaagacttctcctgcgaatagcctgggaaggaataaagga 3840
Db 3781 aaggaagcgttctgggaaacaaagacttctcctgcgaatagcctgggaaggaataaagga 3840
QY 3841 tagagtgttaaaataaaaaaa 3871
Db 3841 tagagtgttaaaataaaaaaa 3871

Db 901 acctcaacccagttatagagagtcacggttgagaaatttgytgcatacatcatt 960
QY 961 acctgaagtttgctcttataaagttaacgacttgaaagaattgagccctctgagctca 1020
Db 961 acctgaagtttgctcttataaagttgaacgacttgaaagaattgagccctctgagctca 1020
QY 1021 gctatctgctctcttgatcagaatgacaaagttcatgacagaaacccgtgtattacagat 1080
Db 1021 gctatctgctctcttgatcagaatgacaaagttcatgacagaaacccgtgtattacagat 1080
QY 1081 acccaaggagacacttggggcctctcgatlgagacacttccagcccaagacctgagcaagttc 1140
Db 1081 acccaaggagacacttggggcctctcgatlgagacacttccagcccaagacctgagcaagttc 1140
QY 1141 agtctttaaagtgaacacacttccagaagagctgtatgacttgcctaaggaataataa 1200
Db 1141 agtctttaaagtgaacacacttccagaagagctgtatgacttgcctaaggaataataa 1200
QY 1201 tggatgatagtgaggagagaagttgtgaaatgtgatgacacctctggaacttgaggaga 1260
Db 1201 tggatgatagtgaggagagaagttgtgaaatgtgatgacacctctggaacttgaggaga 1260
QY 1261 ccagctagccccaagacaacccaagaagactcctctctgagcgttcaaggaaacacagatcct 1320
Db 1261 ccagctagccccaagacaacccaagaagactcctctctgagcgttcaaggaaacacagatcct 1320
QY 1321 tggcctttcccaacagcccaagcctgttatacactcaagacctccttaactctccaaa 1380
Db 1321 tggcctttcccaacagcccaagcctgttatacactcaagacctccttaactctccaaa 1380
QY 1381 gttgaaggagaaagcccccgtctctcttaactgcatgcatcaggggtgagccctgcttcct 1440
Db 1381 gttgaaggagaaagcccccgtctctcttaactgcatgcatcaggggtgagccctgcttcct 1440
QY 1441 atcttcaacactgacacactcaatgtttacacactaactcttctcaactttttttgaga tga 1500
Db 1441 atcttcaacactgacacactcaatgtttacacactaactcttctcaactttttttgaga tga 1500
QY 1501 gttctgctctcttggcccaagcttggaatgtgcaatggcagcttctcaagctcaactgcaacctcc 1560
Db 1501 gttctgctctcttggcccaagcttggaatgtgcaatggcagcttctcaagctcaactgcaacctcc 1560
QY 1561 gctctctgggttcaagaacttctgctgcatcagccctccagagacttgagattacagaca 1620
Db 1561 gctctctgggttcaagaacttctgctgcatcagccctccagagacttgagattacagaca 1620
QY 1621 tgtgccaacagcccgagctaaattgtattcttaagtagagacggggttttgcacatgttg 1680
Db 1621 tgtgccaacagcccgagctaaattgtattcttaagtagagacggggttttgcacatgttg 1680
QY 1681 ccaggtgtgtctcgaactcttgactcagaatgatacactgtgcttggtcctccacaagttgc 1740
Db 1681 ccaggtgtgtctcgaactcttgactcagaatgatacactgtgcttggtcctccacaagttgc 1740
QY 1741 tgggattacaagcgltgagccacatgcccggcctctcttctcaaccttacaacctgtctct 1800
Db 1741 tgggattacaagcgltgagccacatgcccggcctctcttctcaaccttacaacctgtctct 1800
QY 1801 tatctcacaactgttttacaacacttcaacctgtctcttctcaactgttcaactgtctct 1860
Db 1801 tatctcacaactgttttacaacacttcaacctgtctcttctcaactgttcaactgtctct 1860
QY 1861 ccacatgtatagctgctcttcttacaacttgggttggtaaggagagcttctctgtgctg 1920
Db 1861 ccacatgtatagctgctcttcttacaacttgggttggtaaggagagcttctctgtgctg 1920
QY 1921 tttttttgttttcccaagaanaatcaglatattttttaaataagaanaaatcctccagaa 1980
Db 1921 tttttttgttttcccaagaanaatcaglatattttttaaataagaanaaatcctccagaa 1980
QY 1981 gatgataaatgtgaaacactccttggctattgtcttccagatttaagtcctcttc 2040
Db 1981 gatgataaatgtgaaacactccttggctattgtcttccagatttaagtcctcttc 2040

QY 2041 tcccacccgggaagaatgttggaagacataaggctaaatttctccagccctcacaaatgttc 2100
Db 2041 tcccacccgggaagaatgttggaagacataaggctaaatttctccagccctcacaaatgttc 2100
QY 2101 ttcaactgtgtcactgttaccacttctagcacccacttgaaaaaacaagttgtagagag 2160
Db 2101 ttcaactgtgtcactgttaccacttctagcacccacttgaaaaaacaagttgtagagag 2160
QY 2161 tgtagagtgacagaatgttggtcttttgcccaacttgcacatcccaaaattacaacggttgg 2220
Db 2161 tgtagagtgacagaatgttggtcttttgcccaacttgcacatcccaaaattacaacggttgg 2220
QY 2221 ccgactccatttaggacaatgtcttagtataagttcccgagttctggaagaagaagagc 2280
Db 2221 ccgactccatttaggacaatgtcttagtataagttcccgagttctggaagaagaagagc 2280
QY 2281 cagagctgtcacttcatcattccttctcaagtaataattattgagtaactgtgtgc 2340
Db 2281 cagagctgtcacttcatcattccttctcaagtaataattattgagtaactgtgtgc 2340
QY 2341 taaggcatgacacttggaactagagatacttccacagaataacagggaaagtccctgtgct 2400
Db 2341 taaggcatgacacttggaactagagatacttccacagaataacagggaaagtccctgtgct 2400
QY 2401 catgagagcttaacttcttaacagggagaagaagatagccaactataggaataataataac 2460
Db 2401 catgagagcttaacttcttaacagggagaagaagatagccaactataggaataataataac 2460
QY 2461 aaggtatcatgtatgataattgtctgtgagaaanaataaagccgggagaggttaagaa 2520
Db 2461 aaggtatcatgtatgataattgtctgtgagaaanaataaagccgggagaggttaagaa 2520
QY 2521 tccctgagatagagctgcagttttaaactgggcccactcagtgagaaatgtacgttgcagaa 2580
Db 2521 tccctgagatagagctgcagttttaaactgggcccactcagtgagaaatgtacgttgcagaa 2580
QY 2581 gacgttagaggaaatgagatccctgtgacaagagcattcccgagaggaacaaatgtgcactg 2640
Db 2581 gacgttagaggaaatgagatccctgtgacaagagcattcccgagaggaacaaatgtgcactg 2640
QY 2641 ccccaagagtagaactgtcctcagctggttcaagaaagacagggagacccaagacagctgc 2700
Db 2641 ccccaagagtagaactgtcctcagctggttcaagaaagacagggagacccaagacagctgc 2700
QY 2701 tgggcaagggttagaatgtgaagagagagcggtctgggagagacaggtgtgtgagggccttgg 2760
Db 2701 tgggcaagggttagaatgtgaagagagagcggtctgggagagacaggtgtgtgagggccttgg 2760
QY 2761 ctctctgctaagttagatgtgggaacacttgtagaggttttgaacagagagaggtgcttgatgat 2820
Db 2761 ctctctgctaagttagatgtgggaacacttgtagaggttttgaacagagagaggtgcttgatgat 2820
QY 2821 ttataatttgcagaaggttcaacttctagcttgcataatgttgaaaaactttagtggacaagggc 2880
Db 2821 ttataatttgcagaaggttcaacttctagcttgcataatgttgaaaaactttagtggacaagggc 2880
QY 2881 agaaaggaagaggaagaaactgtttaggaagctactgcagaagttccagagcttggcctgggc 2940
Db 2881 agaaaggaagaggaagaaactgtttaggaagctactgcagaagttccagagcttggcctgggc 2940
QY 2941 ccaagcaaacagcagttgttcaaatatctagattatttttgaagaagccaaatagatattgc 3000
Db 2941 ccaagcaaacagcagttgttcaaatatctagattatttttgaagaagccaaatagatattgc 3000
QY 3001 tgaagatttgaatgttgagtgtaagagaaggaagttlaatgatagtacatlaagtttttg 3060
Db 3001 tgaagatttgaatgttgagtgtaagagaaggaagttlaatgatagtacatlaagtttttg 3060
QY 3061 gctctgaatagcaggaagaatgtgaagttacacagttactgaaatlaggaaagatgtggctgggt 3120
Db 3061 gctctgaatagcaggaagaatgtgaagttacacagttactgaaatlaggaaagatgtggctgggt 3120

Db	2041	cccccccccgggaaaagatctggtcggaaagacataagctcaaatcttctccagcctcccaatgctc	2100
Qy	2101	ttcaacttgcctgcctctgttaccacatctccagccaccactgtgaaaaaacaagtttagttagaag	2160
Db	2101	ttcaacttgcctgcctctgttaccacatctccagccaccactgtgaaaaaacaagtttagttagaag	2160
Qy	2161	tgttagatctgcagaaaatgtgtgctcttttgcgccactttgcatctcccaaatatacaagcttgc	2220
Db	2161	tgttagatctgcagaaaatgtgtgctcttttgcgccactttgcatctcccaaatatacaagcttgc	2220
Qy	2221	ccgattcccatcttgggagcaatgtcttggtataagctcccgagtttgaaaaggaagaagc	2280
Db	2221	ccgattcccatcttgggagcaatgtcttggtataagctcccgagtttgaaaaggaagaagc	2280
Qy	2281	cagaagctctcagcttccatctccatctcttccagtgaaatatttatgtagatccactatgtgc	2340
Db	2281	cagaagctctcagcttccatctccatctcttccagtgaaatatttatgtagatccactatgtgc	2340
Qy	2341	taagcatctgacccttgggaaactagagatacttcacagaaataacagggaaaagttccctgtgct	2400
Db	2341	taagcatctgacccttgggaaactagagatacttcacagaaataacagggaaaagttccctgtgct	2400
Qy	2401	catgtgagcttaccatcttccacagggagaaagagataagccataacataagtaataatataac	2460
Db	2401	catgtgagcttaccatcttccacagggagaaagagataagccataacataagtaataatataac	2460
Qy	2461	aaggtatcatcttagtctgataatctgtcttgcgagaaaataaagcagggtggggagtagaagaa	2520
Db	2461	aaggtatcatcttagtctgataatctgtcttgcgagaaaataaagcagggtggggagtagaagaa	2520
Qy	2521	tcctctgagagatgagagctctgcagctttttaaatctgggagccctacatggtatgtgacgttgcaga	2580
Db	2521	tcctctgagagatgagagctctgcagctttttaaatctgggagccctacatggtatgtgacgttgcaga	2580
Qy	2581	gaagcttggggaagtggtgactcctctgggacaagagcatctccagcagagagaaacaagatgtgcactg	2640
Db	2581	gaagcttggggaagtggtgactcctctgggacaagagcatctccagcagagagaaacaagatgtgcactg	2640
Qy	2641	cccccaagctggaactctgtgctctacgtctggtctcagaaaagacacaggagagacaaagcagaagctgcg	2700
Db	2641	cccccaagctggaactctgtgctctacgtctggtctcagaaaagacacaggagagacaaagcagaagctgcg	2700
Qy	2701	tggtgcagaggtgtagaaatgaaaagagagagcggtctggtgagagagacaggtctgtgtgagggcccttgcg	2760
Db	2701	tggtgcagaggtgtagaaatgaaaagagagagcggtctggtgagagagacaggtctgtgtgagggcccttgcg	2760
Qy	2761	cttctgtcgaagtgtgagatctgggaaaccaactgtgagaggtttgaaacagagagatgtgccttatgcat	2820
Db	2761	cttctgtcgaagtgtgagatctgggaaaccaactgtgagaggtttgaaacagagagatgtgccttatgcat	2820
Qy	2821	ttatatcttggaaaaggtgcatctctctacgtctgtaaatatctgtgaaaacactttagtgcgacaagggc	2880
Db	2821	ttatatcttggaaaaggtgcatctctctacgtctgtaaatatctgtgaaaacactttagtgcgacaagggc	2880
Qy	2881	agaaagagagaggtggaagaccctgtttaggaaagcctaactgcaaaagttccagggccttgggccttgcc	2940
Db	2881	agaaagagagaggtggaagaccctgtttaggaaagcctaactgcaaaagttccagggccttgggccttgcc	2940
Qy	2941	cacagcacaagcagctggtctcaaatatactataatttatcttggaaaagggccaataatagatctgc	3000
Db	2941	cacagcacaagcagctggtctcaaatatactataatttatcttggaaaagggccaataatagatctgc	3000
Qy	3001	tgaagagctttaaattgttgtagatgttaaggaaggaagaaattaaatgaatacaatttaagtttttg	3060
Db	3001	tgaagagctttaaattgttgtagatgttaaggaaggaagaaattaaatgaatacaatttaagtttttg	3060
Qy	3061	gctcgttaataagaaagaaatctgagctttaaaccagttactcgtgaataatcgtggaagatctgcgtgcgt	3120
Db	3061	gctcgttaataagaaagaaatctgagctttaaaccagttactcgtgaataatcgtggaagatctgcgtgcgt	3120
Qy	3121	aagtaaggaatttgcgtgcacaagcagctgctgtgtgtgttgaaatgtaggaagttcctgcgtcga	3180
Db	3121	aagtaaggaatttgcgtgcacaagcagctgctgtgtgtgttgaaatgtaggaagttcctgcgtcga	3180

Db	3121	aagtaaggaatttggtgcgaagcagcgctgctctgtgttgtaagggatgtctgtcgtca	3180
Qy	3181	aatcaaatgtgaagatctctcctaaggtcaggtctcgcagcagaagctcgsagacaggatctcgaa	3240
Db	3181	aatcaaatgtgaagatctctcctaaggtcaggtctcgcagcagaagctcgsagacaggatctcgaa	3240
Qy	3241	tgcacttggttatgtttgttgggggtggtcctctaagaagaacctgtgaaagccttatacagtc	3300
Db	3241	tgcacttggttatgtttgttgggggtggtcctctaagaagaacctgtgaaagccttatacagtc	3300
Qy	3301	atttttgcgtctgaagaagtctctcttggaagtgtgtggtacatttgaaagcgaagtgcattca	3360
Db	3301	atttttgcgtctgaagaagtctctcttggaagtgtgtggtacatttgaaagcgaagtgcattca	3360
Qy	3361	gtctgaaggcgaagctctctctgaaagaaggtctgttaagcattctgcaagctacacatgcgtgtag	3420
Db	3361	gtctgaaggcgaagctctctctgaaagaaggtctgttaagcattctgcaagctacacatgcgtgtag	3420
Qy	3421	tgtgttgggggttgggggtctcttggcacgtgctgtgtgaaggatctctgcaaggcaccaca	3480
Db	3421	tgtgttgggggttgggggtctcttggcacgtgctgtgtgaaggatctctgcaaggcaccaca	3480
Qy	3481	ggccccctacttgaaacataagcagctgacatgtgacatttaagaccagtcagcgtggaggggc	3540
Db	3481	ggccccctacttgaaacataagcagctgacatgtgacatttaagaccagtcagcgtggaggggc	3540
Qy	3541	cactgaagatctctctctgaattactctaagaagaacacagaagaagcagctgaatggagcc	3600
Db	3541	cactgaagatctctctctgaattactctaagaagaacacagaagaagcagctgaatggagcc	3600
Qy	3601	tgtggtctctctgggaaatvggaaatcagccaaagagctggaagaaggttaacctaaagtc	3660
Db	3601	tgtggtctctctgggaaatvggaaatcagccaaagagctggaagaaggttaacctaaagtc	3660
Qy	3661	agaagaagaacccaagagagtggtgtgttcttgaaagtctgagctcttattcaaccccatctc	3720
Db	3661	agaagaagaacccaagagagtggtgtgttcttgaaagtctgagctcttattcaaccccatctc	3720
Qy	3721	ccttctccaaataaagcacactgtgtagtttgggccctccaggggttgaaagcgaaggagga	3780
Db	3721	ccttctccaaataaagcacactgtgtagtttgggccctccaggggttgaaagcgaaggagga	3780
Qy	3781	aagcagcagcgttctgggaaacagaactctctctgcaatagccttgggaaggaataaagga	3840
Db	3781	aagcagcagcgttctgggaaacagaactctctctgcaatagccttgggaaggaataaagga	3840
Qy	3841	tagagtgcttaaatataaaaaaataaaaaa	3871
Db	3841	tagagtgcttaaatataaaaaaataaaaaa	3871
RESULT 4			
AAFI6225			
ID	AAFI6225 standard; cDNA; 2549 BP.		
AC	AAFI6225:		
XX			
DT	13-MAR-2001 (first entry)		
XX			
DE	Human prostate cancer antigen nucleotide sequence SEQ ID NO:660.		
XX			
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;		
KW	neuroprotective; cytosolic; prostatic; cardiovascular; immunomodulatory; muscular;		
KW	vulnerarity; gastrointestinal; nephrotropic; antiinfective; gynaecological;		
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;		
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;		
XX	wound; infectious disease; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200055174-A1.		
XX			
PD	21-SEP-2000.		

XX	PF	08-MAR-2000; 2000WO-US05988.
XX	PR	12-MAR-1999; 99US-0124270.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	PA	(ROSE/) ROSEN C A.
XX	P1	Rosen CA, Ruben SM;
XX	DR	WPI: 2000-587513/55.
XX	DR	P-PSDB; AAB57022.
PT	PT	Prostate cancer associated gene sequences, referred to as prostate
PT	PT	cancer antigens, useful for treatment, prevention, and diagnosis of
PT	PT	disorders such as prostate cancer -
XX	XX	Claim 1; Page 1097-1098; 2338pp; English.
XX	XX	AAP15566 to AAP16505 encode the human prostate cancer associated
CC	CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC	CC	The prostate cancer antigens can have neuroprotective, cytoskeletal,
CC	CC	cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC	CC	nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC	CC	and can be used in gene therapy. The prostate cancer antigen
CC	CC	polynucleotides may be used for detection of prostate cancer, chromosome
CC	CC	identification, as chromosome markers, and for numerous other diagnostic
CC	CC	or research purposes. The prostate cancer antigens may be used to treat
CC	CC	disorders such as neural, immune, muscular, reproductive,
CC	CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC	CC	disorders, wounds, and infectious diseases. AAP16506 to AAP16514 to
CC	CC	AAB57303 represent sequences used in the exemplification of the present
CC	CC	invention.
SQ	XX	Sequence 2549 BP; 667 A; 519 C; 683 G; 673 T; 7 other;
Query Match	53.0%; Score 2051.2; DB 21; Length 2549;	
Best Local Similarity	99.2%; Pred. No. 0;	
Matches 2077; Conservative	6; Mismatches	4; Indels
		7; Gaps
YY	1763	catgcggcctcttccacacttaacacgttccttatccatacatcttttcaaa 1822
Dd	206	ctcgccggcctcttccacacttaacacgttccttatccatacatcttttcaaa 265
YY	1823	ccttaaccctgtctcctcatagttcacacttgcttcccgaatgtcataagtgccttc 1882
Dd	266	ccttaaccctgtctcctcatagttcacacttgcttcccgaatgtcataagtgccttc 325
YY	1883	ttaaccttggittgaaggcagctctctcgttgttttttggttttccccgaanaa 1942
Dd	326	ttacaatttggittgaaggcagctctctcgttgttttttggttttccccgaanaa 385
YY	1943	tcagtatattttttaaataagaanaaacattccctgaagaatgataaatgtgaanaacctcc 2002
Dd	386	tcagtatattttttaaataagaanaaacattccctgaagaatgawattgtggaanaacctcc 445
YY	2003	tttggctatttgcctttccagatttagtctccttccctccacccgggaagaatgtgtg 2062
Dd	446	tttggctatttgcctttccagatttagtctccttccctccacccgggaagaatgtgtg 505
YY	2063	gaagcacaaggctaatttccagccccaacabtgcttcaactgtgtcgtactgtacc 2122
Dd	506	gaagcacaaggctaatttccagccccaacabtgcttcaactgtgtcgtactgtacc 565
YY	2123	aattctagcaccacatgaanaaacaagtgtgagtagagagtgtagagtgacgaatgtgct 2182
Dd	566	aattctagcaccacatgaanaaacaagtgtgagtagagtagagtgtagaagaatgtgct 625
YY	2183	tttgcaccaattgtcattccaaaataatacaacggtgttgccgatcccatgttgsagaacatg 2242
Dd	626	tttgcaccaattgtcattccaaaataatacaacggtgttgccgatcccatgttgsagaacatg 685

QY	2243	cttggttaaggtctccgaggttggaaaagaaagaaagccagagctgtctagttaattca	2302
Db	686	cttagttaaggtctccgaggttggaaaagaaagaaagccagagctgtctagttaattca	745
QY	2303	ttcttcaagtaaatattattgtggttacctctgtgtctaggcatgtgacctg-----g	2356
Db	746	ttcttccagtaaatattattgtggttacctctgtgtctaggcatgtgacctggaactag	805
QY	2357	aactagagatacttcacagataacaaagggaaagttccctgtgtctcaitgagcttcaatc	2410
Db	806	aactagagatacttcacagataacaaagggaaagttccctgtgtctcaitgagcttcaatc	865
QY	2417	tacagggagaaagagatagcccaatacacaiaagaaataataatacaagatcacatgtag	2476
Db	866	tacagggagaaagagatagcccaatacacaiaagaaataataatacaagatcacatgtag	925
QY	2477	ataatgtgtgtgggaaaaaaataaagcaggggagggaggtgaagaaatcccttgagataagagct	2538
Db	926	ataatgtgtgtgggaaaaaaataaagcaggggagggaggtgaagaaatcccttgagataagagct	985
QY	2537	gcagtttcaaaatgtgggctccactgtggagatgtgacgtcttgacagagacgtctaaaggaggtg	2596
Db	986	gcagtttcaaaatgtgggctccactgtggagatgtgacgtcttgacagagacgtctaaaggaggtg	10454
QY	2597	atcctgtgacaagagcatccacagcgacagagagaaacaagatgtgcactgtgcccaagtgtgaact	2655
Db	1046	atcctkgacaaagagcmcttcccgagctcagagaaacaagatgtgcactgtgcccaagtgtgaact	1105
QY	2657	tgctctaaaggtgttcaaggaaagagcagagggaaaccaaagcaagatctgtgtgagaggtgtagaat	2711
Db	1106	tgctctaaaggtgttcaaggaaagagcagagggaaaccaaagcaagatctgtgtgagaggtgtagaat	1165
QY	2717	ggaagagagagcgctgtggggagagacaggtgtgtgagagggctcttgctctcgtccaaagtga	2776
Db	1166	ggaagagagagcgctgtggggagagacaggtgtgtgagagggctcttgctctcgtccaaagtga	1225
QY	2777	tgggaaaccatgtgaggggttggaaacagagagatgctcttgatattatatttccaaagg	2833
Db	1226	tgggaaaccatgtgaggggttggaaacagagagatgctcttgatattatatttccaaagg	1285
QY	2837	tcatctcagctgcgaatacttctgaaaaactttagtgaaacaaggccaaagaaagagagagag	2896
Db	1286	tcatctcagctgcgaatacttctgaaaaactttagtgaaacaaggccaaagaaagagagagag	1345
QY	2897	accgtgttagagagcttaactgtgaaaaggttccaaagctctgggctctgggcccacagcaacagatg	2956
Db	1346	accgtgttagagagcttaactgtgaaaaggttccaaagctctgggctctgggcccacagcaacagatg	1405
QY	2957	gtcaaatatctagaattattatttgaaaaagccaataagatatgtgtgtgagaggttgaatgtg	3018
Db	1406	gtcaaatatctagaattattatttgaaaaagccaataagatatgtgtgtgagaggttgaatgtg	1465
QY	3017	gaggtgttaagagaaagaaagatttaatgatgacattaaagtttttggcctgtgataagagaa	3076
Db	1466	gaggtgttaagagaaagaaagatttaatgatgacattaaagtttttggcctgtgataagagaa	1523
QY	3077	agatgtgaattcccgcttaactgaaataagggaaagatatgtgctgttgtaagtaagtaattgtg	3138
Db	1526	agatgtgaattcccgcttaactgaaataagggaaagatatgtgctgttgtaagtaagtaattgtg	1585
QY	3137	gcaaaagcagagctgtctgtgtgtgtgaaatgtggaaggttctgtgctgcaaaatacaagtgtgaa-t	3195
Db	1586	gcaaaagcagagctgtctgtgtgtgtgaaatgtggaaggttctgtgctgcaaaatacaagtgtgaa-t	16454
QY	3196	tctctcaggtcaggtctctgcagccagagctcggagacagagatctggaatgcactgtgttat	3255
Db	1646	tctctcaggtcaggtctctgcagccagagctcggagacagagatctggaatgcactgtgttat	1705
QY	3256	gttggggggtgtcttcacaaagaaacccgttgaaaaccccttatacgttaattatgtgtgtga	3315
Db	1706	gttggggggtgtcttcacaaagaaacccgttgaaaaccccttatacgttaattatgtgtgtga	1765
QY	3316	gaagtctctcgtggaggtgtgtgtgtaacatttgaaagggagatctacagtttgagggcagagct	3375

Db	1766	gaagttcttcgaggagctgtggtgtacatttgaaagcgaaagtacttcagttcgagggcaagcc	18235
OY	3376	ctggaagaagagcgtgtagagcatctggcagcagtaacatcgtgtgtaagtgtcttgggggtggg	34355
Db	1826	ctggaagaagagcgtgtagagcatctggcagcagtaacatcagtagtgtgtgtgggggtggg	18855
OY	3436	gtcttcgggcacttgcgtgtgttgaaaggatcttgcagaggcacaacagcgccctactgaa	34355
Db	1886	gtcttcgggcacttgcgtgtgttgaaaggatcttgcagaggcacaacagcgccctactgaa	19455
OY	3496	ccatcacatctacgtgtgcatattaaagccatgcagcttggaaggcgccactggaatgtctc	35555
Db	1946	ccatcacatctacgtgtgcatattaaagccatgcagcttggaaggcgccactggaatgtctc	20055
OY	3556	tgaattactctgagaagcaacagaaaaagccatgtgataggacctctggcctctctggga	36155
Db	2006	tgaattactctgagaagcaacagaaaaagccatgtgataggacctctggcctctctggga	20655
OY	3616	aatgggaatctcagccaaagagactgtgaaagaaggtttacctaaagtcagaagaaaaccaagag	36755
Db	2066	aatgggaatctcagccaaagagactgtgaaagaaggtttacctaaagtcagaagaaaaccaagag	21255
OY	3676	agttgtgtgtctcggaaagcgtgagccttcttattcaacttaacttccctcccaataag	37355
Db	2126	agttgtgtgtctcggaaagcgtgagccttcttattcaacttaacttccctcccaataag	21855
OY	3736	ccactgtgtgaagtgtggccctcccaagggttgaaagcgaaagaggaagaaagcacaagcttgg	37955
Db	2186	ccactgtgtgaagtgtggccctcccaagggttgaaagcgaaagaggaagaaagcacaagcttgg	22455
OY	3796	ggaacaaagactttctcgtgaatagaccctgggaaggaataaaagataagtgct	3849
Db	2246	ggaacaaagactttctcgtgaatagaccctgggaaggaataaaagataagtgct	2299

RESULT	5
AAH14525	
ID	AAH14525 standard; cDNA; 1961 BP.
XX	
AC	AAH14525;
XX	
DT	26-JUN-2001 (first entry)
XX	
DE	Human cDNA sequence SEQ ID NO:12066.
XX	
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss
XX	
OS	Homo sapiens.
XX	
PN	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000EP-0116126.
XX	
PR	29-JUL-1999; 99JP-0248036.
PR	27-AUG-1999; 99JP-0300253.
PR	11-JAN-2000; 2000JP-0118776.
PR	02-MAY-2000; 2000JP-0183767.
PR	09-JUN-2000; 2000JP-0241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	Isihli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	
DR	WPI; 2001-318749/34.
XX	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602
PT	full-length cDNAs defined in the specification, and for the detection
PT	and/or diagnosis of the abnormality of the proteins encoded by the
PT	full-length cDNAs -

XX Claim 8; SEQ ID 12066; 2537pp + CD ROM; English.
PS

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1961 BP; 417 A; 579 C; 481 G; 484 T; 0 other;

Query Match	50.58;	Score 1955.4;	DB 22;	Length 1961;
-------------	--------	---------------	--------	--------------

Best Local Similarity 99.78; Pred. No. 0;

[illegible]

OY	624	taaagaagccctgcttgctgcgcagagacttaattaaagacccctggaaccagatcatgaa	683
Db	601	tataagaagccgcgcctgctgcgcagagacttaattaaagacccctggaaccagatcatgaa	660
OY	684	aagcctgctcatccagacagatgctgcgagcatatacaacggtgagaaactggaacatccatcca	743
Db	661	agccctgtcaatccagagacagatgctgcgagcatatacaacggtgagaaactggaacatccatcca	720
OY	744	gacatgagcctgagccctctccgaactcttcaaaaagcctttacagagtgtctcgcaagcttgc	803
Db	721	gacatgagcctgagccctctccgaactcttcaaaaagcctttacagagtgtctcgcaagcttgc	780
OY	804	gagagcttccagaggagatccaagagcttccaagagatttactacttccatagcagatcatat	863
Db	781	gagagcttccagaggagatccaagagcttccaagagatttactacttccatagcagatcatat	840
OY	864	gtaagaagtctctggaatgcaaaataaagctgtaagaagaaacctcaaccacattatagagagc	923
Db	841	gtagaagttctctggaatgcaaaataaagctgtaagaagaaacctcaaccacattatagagagc	900
OY	924	tatccggtctgagaaaatttgtgcttaacatgtaatcattactgtcaggttgcctatataga	983
Db	901	tatccggtctgagaaaatttgtgcttaacatgtaatcattactgtcaggttgcctatataga	960
OY	984	ttgagaagaccttgagaagaatttgagacccctctgtgcacgttaagtatactgtcctttgataagaat	1043
Db	961	ttgagaagaccttgagaagaatttgagacccctctgtgcacgttaagtatactgtcctttgataagaat	1020
OY	1044	gacaagaagctcatgcaagacagaaacctgtggtatattacaagtacaagaaggaaaccttggagccct	1103
Db	1021	gacaagaagctcatgcaagacagaaacctgtggtatattacaagtacaagaaggaaaccttggagccct	1080
OY	1104	tcgagatgagacattccagaccagaaaccttgaaagcagtttcagttctttaaattgagccacact	1163
Db	1081	tcgagatgagacattccagaccagaaaccttgaaagcagtttcagttctttaaattgagccacact	1140
OY	1164	cagaagaagagcctgtatgaacttctgtaagaagaataataatgagatgatatgagaggagaagt	1223
Db	1141	cagaagaagagcctgtatgaacttctgtaagaagaataataatgagatgatatgagaggagaagt	1200
OY	1224	gtggaataatbtggaatgagccctctggaaactggagaggagacaaggttaagcccaagcaaacaa	1283
Db	1201	gtggaataatbtggaatgagccctctggaaactggagaggagacaaggttaagcccaagcaaacaa	1260
OY	1284	gagagcttccctctgtgcgttcaaggaacaacagatattcttgcctttcccaaaagcccgag	1343
Db	1261	gagagcttccctctgtgcgttcaaggaacaacagatattcttgcctttcccaaaagcccgag	1320
OY	1344	ctgtgtgataccccaagagccttctcttactctccaagtgaagaaggagcccccgtctct	1403
Db	1321	ctgtgtgataccccaagagccttctcttactctccaagtgaagaaggagcccccgtctct	1380
OY	1404	ctaaactgcatgtatatagaagggtgtagagcttgccttctactccaactgccaactgccaactatg	1463
Db	1381	ctaaactgcatgtatatagaagggtgtagagcttgccttctactccaactgccaactgccaactatg	1440
OY	1464	ttcaacacctaacttcttccaaactttttttgagatgagagctgcgctctgtcccaagagtg	1523
Db	1441	ttcaacacctaacttcttccaaactttttttgagatgagagctgcgctctgtcccaagagtg	1500
OY	1524	gagtgcaatgcaagcgttctccagctcaactgcaaaccttccgcctcttgggtttcaagaacatct	1583
Db	1501	gagtgcaatgcaagcgttctccagctcaactgcaaaccttccgcctcttgggtttcaagaacatct	1560
OY	1584	gctgcatcaagccctcccgagttactgtggtatataagcattgccaagtccaaccagcccggttaatt	1643
Db	1561	gctgcatcaagccctcccgagttactgtggtatataagcattgccaagtccaaccagcccggttaatt	1620
OY	1644	tttgtatttttaagtagagaacgggtttttgtgcacatttgccaaggtctgtctcgaacactttgaa	1703
Db	1621	tttgtatttttaagtagagaacgggtttttgtgcacatttgccaaggtctgtctcgaacactttgaa	1680
OY	1704	cttcaagaatgatacatgctgtgccttccccaagatgcttggtatataaagcgttgagcaacc	1763

Db	1681	cttcagatgatcatcctgcttgcgtccccaagctgggattacaagcgtgagccac	1740
Qy	1764	atgcccgcgtcttcttcacaccttaccacgtctctctatcctccacacgttttccac	1833
Db	1741	atgcccgcgtcttcttcacaccttaccacgtctctctatcctccacacgttttccac	1800
Qy	1824	cttcatcctgtcttcttcacacgttcttcccatggtacatagctgccttct	1883
Db	1801	cttcatcctgtcttcttcacacgttcttcccatggtacatagctgccttct	1860
Qy	1884	taccatttgggttaagagcgactcttcctgcgtctgtttttgtttttccagaat	1943
Db	1861	taccatttgggttaagagcgactcttcctgcgtctgtttttgtttttccagaat	1920
Qy	1944	cagtatatttttaataagaanaacatcctcagaagatg	1984
Db	1921	cagtatatttttaataagaanaacatcctcagaagatg	1961
RESULT	6		
AAH31138			
ID	AAH31138	standard; cDNA; 1586 BP.	
XX	AAH31138;		
AC	AAH31138;		
XX			
DT	27-JUL-2001	(first entry)	
XX			
DE	Human colon cancer cell line Km12L4-A cDNA library derived seq#1072.		
XX			
KW	Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;		
XX	detection; colon cancer cell line Km12L4-A; ss.		
OS	Homo sapiens.		
XX			
PN	WO200018916-A2.		
XX			
PD	06-APR-2000.		
XX			
PF	23-SEP-1999;	99WO-US22226.	
XX			
PR	28-SEP-1998;	98US-0102161.	
PR	28-SEP-1998;	98US-0102180.	
PR	29-SEP-1998;	98US-0102380.	
PR	08-OCT-1998;	98US-0103815.	
PR	27-OCT-1998;	98US-0105877.	
XX			
PA	(CHIR) CHIRON CORP.		
XX	(HYSE-) HYSEQ INC.		
PI	Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;		
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;		
PI	Lamson G, Drmanac R, Cirvenjakov R, Dickson M, Drmanac S, Labat I;		
PI	Leshkowitz D, Kita D, Garcia V, Jones IW, Stache-Crain B;		
XX			
DR	WPI; 2000-293155/25.		
XX			
PT	Polynucleotide library comprising 1079 defined sequences, useful in		
XX	the form of an array to detect cancer or susceptibility to cancer -		
PS	Claim 1; Page 498-499; 502pp; English.		
XX			
CC	The present invention describes a library of polynucleotides comprising		
CC	1079 nucleotide sequences (given in AAH31067 to AAH31145). Also described		
CC	are: (1) an isolated polynucleotide (I) having at least 90% identity to		
CC	one of the 1079 sequences; (2) a recombinant host cell containing (1);		
CC	(3) an isolated polypeptide (II) encoded by (1); (4) an antibody that		
CC	specifically binds to (II); (5) a vector comprising (I); and (6) a method		
CC	of detecting differentially expressed genes correlated with a cancerous		
CC	state of a mammalian cell comprising detecting a gene product encoded by		
CC	65 of the 1079 sequences given in the specification. The polynucleotides		
CC	are used to monitor patients having (or susceptible) to cancer to detect		
CC	potentially malignant events at a molecular level before they are		

CC detectable at a gross morphological level. The polynucleotides are also
 CC useful for monitoring the efficacy of various therapies and preventive
 CC interventions. Polynucleotide probes based on the disclosed sequences
 CC are useful for chromosome mapping and detection of transcription levels.
 CC The 1079 polynucleotide sequences were derived from a human colon cancer
 CC cell line Km12L4-A cDNA library.

XX
 SQ Sequence 1586 BP; 381 A; 406 C; 340 G; 453 T; 6 other;

Query Match 38.1%; Score 1474.4; DB 21; Length 1586;
 Best Local Similarity 98.3%; Pred. No. 1.3e-287;
 Matches 1506; Conservative 5; Mismatches 16; Indels 5; Gaps 2;

QY 559 ccgcgtcacacacaccttactgaagcattccctgtagacgaataatgtagaaggaacatg 618
 Db 1 ccgcgtcacacacaccttactgaagcattccctgtagacgaataatgtagaaggaacatg 60
 QY 619 catattataagagcctctgctgctgcgagactatacatataagacctggaacccaagtcac 678
 Db 61 catattataagagcctctgctgctgcgagactatacatataagacctggaacccaagtcac 120
 QY 679 atgaaagccttataccgagcagtgccgagcatatacaagtgagagaactcgagaaatcca 738
 Db 121 atgaaagccttataccgagcagtgccgagcatatacaagtgagagaactcgagaaatcca 180
 QY 739 tcacagacatgagcgtgcccctcccgactcttcaaaagcctttacagagtgctcgag 798
 Db 181 tcacagacatgagcgtgcccctcccgactcttcaaaagcctttacagagtgctcgag 240
 QY 799 cctcgagaggttcgaaggagatcaaggaactcaagattcttaaccttcataagacagac 858
 Db 241 cctcgagaggttcgaaggagatcaaggaactcaagattcttaaccttcataagacagac 300
 QY 859 attatgtgaagttcttggaatgcaaaatacagtgtagaagaagaacctccacccaagtacag 918
 Db 301 attatgtgaagttcttggaatgcaaaatacagtgtagaagaagaacctccacccaagtacag 360
 QY 919 gaggcataccggttgaagaattctgtgctacatgatacatctactgtgaagttgctcatt 978
 Db 361 gaggcataccggttgaagaattctgtgctacatgatacatctactgtgaagttgctcatt 420
 QY 979 ataagttgaagcactggaagaatcagccccctgtgcagtcagctatctgctcttgatc 1038
 Db 421 ataagttgaagcactggaagaatcagccccctgtgcagtcagctatctgctcttgatc 480
 QY 1039 agaattgaacagctcagcagcagaacctggtgtattaccagttaccacagggagacacttgg 1098
 Db 481 agaattgaacagctcagcagcagaacctggtgtattaccagttaccacagggagacacttgg 540
 QY 1099 gcctctcgatgacacttccagcccaagcactgaagcagttcagttctttaatgtgacca 1158
 Db 541 gcctctcgatgacacttccagcccaagcactgaagcagttcagttctttaatgtgacca 600
 QY 1159 cactccagaagaagcgtgatgacttgcataaggaataataatgtagatgagggag 1218
 Db 601 cactccagaagaagcgtgatgacttgcataaggaataataatgtagatgagggag 660
 QY 1219 aagttgtgaatatgtgatacactctctggaactggaaggaacagcagccacagaacaa 1278
 Db 661 aagttgtgaatatgtgatacactctctggaactggaaggaacagcagccacagaacaa 720
 QY 1279 ccaaaagacacttctctctgtgcttcgaagaaacacagattcttctcctttcccaacagc 1338
 Db 721 ccaaaagacacttctctctgtgcttcgaagaaacacagattcttctcctttcccaacagc 780
 QY 1339 ccagagctgtgatacctcaagagccttcttactctccaaagtgaagaggaagccccc 1398
 Db 781 ccagagctgtgatacctcaagagccttcttactctccaaagtgaagaggaagccccc 840
 QY 1399 tctctctaacgcattgatacaggggtgtagcctgcttctctatcttaacacctgacac 1458
 Db 841 tctctctaacgcattgatacaggggtgtagcctgcttctctatcttaacacctgacac 900

QY 1459 tcaattcacacatactcttccacactttttttagatgagatcgcctctcccca 1518
 Db 901 tcaattcacacatactcttccacactttttttagatgagatcgcctctcccca 960
 QY 1519 ggcctgagatgcaatgagcagcttccagctcaactgcaactccgcctcttggtccaaga 1578
 Db 961 ggcctgagatgcaatgagcagcttccagctcaactgcaactccgcctcttggtccaaga 1020
 QY 1579 attcgtgcatacagcctcccgagttacctggatgatacagagatggtccacagcccgagc 1638
 Db 1021 attcgtgcatacagcctcccgagttacctggatgatacagagatggtccacagcccgagc 1080
 QY 1639 taatttgatttttagtagagagagcgggttttgcattgttgccagcagctggtctcgaact 1698
 Db 1081 taatttgatttttagtagagagagcgggttttgcattgttgccagcagctggtctcgaact 1140
 QY 1659 ctgagctcagatgatacattgcctctgctgctcccaagctgctggatatacagcgctgag 1738
 Db 1141 ctgagctcagatgatacattgcctctgctgctcccaagctgctggatatacagcgctgag 1200
 QY 1759 ccacatgcccggcctctcttccacacttaccacgtctcttatacctacatctgttt 1818
 Db 1201 ccacatgcccggcctctcttccacacttaccacgtctcttatacctacatctgttt 1260
 QY 1819 cacacattcacctcctgtctcttccatgcttcaactgtgtcttcccaatgatacagctgac 1878
 Db 1261 cacacattcacctcctgtctcttccatgcttcaactgtgtcttcccaatgatacagctgac 1320
 QY 1879 ttcttaccatttgggttgaaggcagctcttccctgctgcttttttggttttccag 1938
 Db 1321 ttcttaccatttgggttgaaggcagctcttccctgctgcttttttggttttccag 1380
 QY 1939 aaatcacatattatttttaataaagaanaaacattccagagaatgataatgtgaaac 1998
 Db 1381 aaatcacatattatttttaataaagaanaaacattccagagaatgataatgtgaaac 1440
 QY 1999 ctcccttgcttaattgcttttccag--atttagtctccttctccacacgggaaag 2056
 Db 1441 ctcccttgcttaattgcttttccagatttaagtctccttctccacacgggaaag 1500
 QY 2057 a---tggtggaagacataggctaattctcc 2085
 Db 1501 agatgggtggaagacataagctaaatcttc 1532

RESULT 7

AAH06808
 ID AAH06808 standard; cDNA: 818 BP.

XX AC AAH06808;

XX CT 26-JUN-2001 (first entry)

XX DE

XX Human cDNA clone (5'-primer) SEQ ID NO:3643.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX FN EP1074617-A2.

XX XX

XX PD 07-FEB-2001.

XX PF

XX 28-JUL-2000; 2000EP-0116126.

XX XX

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX XX

XX PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
PS
XX Claim 1: SEQ ID 3643; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95693 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 818 BP; 159 A; 271 C; 236 G; 148 T; 4 other;

[illegible]

Dd	421	gcttcgcgcagctccacggcccaggccggagtgtcgtgcgacttccagcgccgagcccc	480
Oy	504	tacaagttcctgcgtgcttgcttactcaaggcaataatctccccaagccatgcogct	563
Dd	481	tacaagttcctgcgtgcttgcttactcaaggcaataatctccccaagccatgcogct	540
Oy	564	gctcacaccttcttactgaagaatcccgatgagaaatgattgaagngaacatgcatat	623
Dd	541	gctcacaccttcttactgaagaatcccgatgagaaatgattgaagngaacatgcatat	600
Oy	624	tataagagccgctcgtggtgcgcgagactaacattaaagacttggaaaaccaagt--cata	681
Dd	601	tataagagccgctcgtggtgcgcgagactaacattaaagacttggaaaaccaagt--cata	659
Oy	682	aaagcctgttcatccgcagcagltgcgggcatalacaacggttgaagacttggagaacatcalca	741
Dd	660	aaagcctgttcatccgcagcagltgcgggcatalacaacggttgaagacttggagaacatcalca	719
Oy	742	cagacatgagagctggccc--ttccggaacttcttaagcccttt	783
Dd	720	cagacatgagagctggccccttccgaatttcttcaaaagncttt	763
RESULT	8		
ID	AAZ80131/c		
XX	AAZ80131 standard; cDNA: 618 BP.		
AC	AAZ80131;		
XX	07-APR-2000 (first entry)		
DE	Human colon cancer cell line SW480 cDNA clone SEQ ID NO:215.		
KW	Human: gene expression product; diagnosis; tumour; colon cancer; colorectal adenocarcinoma; cell line SW480; cell proliferation; cytosolic; sarcoma; breast cancer; neoplasia; dysplasia; hyperplasia; ds.		
OS	Homo sapiens.		
PN	WO9964576-A2.		
PD	16-DEC-1999.		
PF	09-JUN-1999; 99WO-IB01062.		
PR	10-JUN-1998; 98US-0088801.		
PA	(FARB) BAYER CORP.		
P1	Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE; Carroll E, Catlino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE; Schlegel R;		
WI:	2000-087220/07.		
PT	Novel nucleic acids, used to develop products for the diagnosis and treatment of disorders involving unwanted cell proliferation, particularly cancers, especially colon cancer		
PS	Claim 15; Page 218; 469pp; English.		
CC	AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from the human colorectal adenocarcinoma (colon cancer) cell line SW480. The cDNA clones can be used to generate antisense oligonucleotides which can be used for antisense therapy. Methods and products from the present invention can be used for identifying and/or classifying cancerous cells present in a human tumour, particularly in solid tumours, e.g. carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones can be used for developing agents for the diagnosis and treatment of disorders involving unwanted cell proliferation, such as neoplasia, dysplasia or hyperplasia.		

SO Sequence 618 BP; 157 A; 133 C; 169 G; 136 T; 23 other:

Query Match 12.0%; Score 466.4; DB 21; Length 618;
Best Local Similarity 94.8%; Pred. No. 1.2e-84;
Matches 506; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

Oy 1076 ccagtaacacagaggaacttggagcctcgcgagatgagcacttccagccagactgaac 1135
D 534 CNAGTCCCNNAAGGCCANTGGGCTMTNGATGAGCANTTCCAGCCAGCTTAANAC 475
Oy 1136 agttcagttcttaattgagcca-cactccagaagaagc-gtattgactt-gctaaaga 1192
D 474 AGTTGAGTTCTTAATATGNCACCCACTCCAGANAGACTTGTATGACTTTGGNTAAGA 415
Oy 1193 aaataaagatgatgatgagaggaagttgtggaataatgtagagcccttgaact 1252
D 414 AATATATATGATGATGATGAGGAGAGTGTGSAATATGATGATGATGATGATGAT 355
Oy 1253 ggaagagacagcctagccacagaacccaagaagacttctcttgaggtcagaagaac 1312
D 354 GGAGAGAGACAGTTAGCCACAGCAACCAAGAGACTTCTTGGCGCTCAGGAAACAC 295
Oy 1313 agattcttgccttttcccaagagccaggtgtgtatcctaagagccttcttcttac 1372
D 294 AGATTCTTGTCTTTCCTTCCACAGCCAGCTGTGTATACCTCAGACCTTCTTCTTAC 235
Oy 1373 tctcaaaagttaaaaggaagcccgctctccttaactcatgcatcaggggtgagcctg 1432
D 234 TCTCAAAAGTAAAGGGAAGGCCCGCTCTCTTAATCTCATGTCATCAGGGGTGAGCCTG 175
Oy 1433 ccttctctatctcaacacttgcacactcaatgttcaacactatcttctacactttttt 1492
D 174 CTTTCCATCTTACACACTCCACCTCATCTTACACCTTCTTCTTCTTCTTCTTCTTCT 115
Oy 1493 gaggtgagctctgcctcctcttgccagagctgagtgcaatgacagcttctgaactctg 1552
D 114 GAGATGAGGATCTCGCTCTCTTGGCCAGGCTGGAAGTGCACGTTCTCAGCTCATCTG 55
Oy 1553 caactcgcgcctcttggttcaagaacttctgcatcagcctccagatcacc 1606
D 54 CAACCTCGCGCTTGGGTTCAAGCAATTCGTGATCAGACCTCCGCAAGTAC 1

RESULT 9
AAH30585
ID AAH30585 standard; cDNA; 408 BP.
XX
AC AAH30585;
XX
DT 27-JUL-2001 (first entry)
XX
DE Human colon cancer cell line Km12L4-A cDNA library derived sequence #519.
XX
KW Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Km12L4-A; ss.
XX
OS Homo sapiens.
XX
PN MO200018916-A2.
PD 06-APR-2000.
XX
PF 23-SEP-1999; 99WO-US22226.
XX
PR 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105877.
XX
PA (CHIR) CHIRON CORP.
PA (HSE-) HYSED INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Cravenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LM, Stache-Crain B;
XX
DR WPI: 2000-293155/25.
XX
PT Polynucleotide library comprising 1079 defined sequences, useful in
XX the form of an array to detect cancer or susceptibility to cancer -
XX
PS Claim 1; Page 333; 502pp; English.

CC The present invention describes a library of polynucleotides comprising
CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC one of the 1079 sequences; (2) a recombinant host cell containing (1);
CC (3) an isolated polypeptide (II) encoded by (1); (4) an antibody that
CC specifically binds to (II); (5) a vector comprising (I); and (6) a method
CC of detecting differentially expressed genes correlating with a cancerous
CC state of a mammalian cell comprising detecting a gene product encoded by
CC 65 of the 1079 sequences given in the specification. The polynucleotides
CC are used to monitor patients having (or susceptible) to cancer to detect
CC potentially malignant events at a molecular level before they are
CC detectable at a gross morphological level. The polynucleotides are also
CC useful for monitoring the efficacy of various therapies and preventive
CC interventions. Polynucleotide probes based on the disclosed sequences
CC are useful for chromosome mapping and detection of transcription levels.
CC The 1079 polynucleotide sequences were derived from a human colon cancer
XX cell line Km12L4-A cDNA library.

SO Sequence 408 BP; 117 A; 96 C; 98 G; 97 T; 0 other:

Query Match 10.5%; Score 406.4; DB 21; Length 408;
Best Local Similarity 99.8%; Pred. No. 1.3e-72;
Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 559 ccgctgtcacaccttctactgaagcctcctgaatgaagaaatgtaagaaggaacatg 618
D 1 ccgctgtcacaccttctactgaagcctcctgaatgaagaaatgtaagaaggaacatg 60
Oy 619 catattataagagcctgctgtgctcgaagactcatataaagacctggaacaaatgcat 678
D 61 catattataagagcctgctgtgctcgaagactcatataaagacctggaacaaatgcat 120
Oy 679 atgaagcctgttcatccgagcagtgcgagatacaacggtggaagactggaagacatcca 738
D 121 atgaagcctgttcatccgagcagtgcgagatacaacggtggaagactggaagacatcca 180
Oy 739 tcacagacatgagctggccctcccgactcttcaagaagccttctacagtgctcgag 798
D 181 tcacagacatgagctggccctcccgactcttcaagaagccttctacagtgctcgag 240
Oy 799 cctgcgaggttccagaggaatcaaggaactcaaggaattctacattccatagcagatc 858
D 241 cctgcgaggttccagaggaatcaaggaactcaaggaattctacattccatagcagatc 300
Oy 859 attatgtagaagctctggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 918
D 301 attatgtagaagctctggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 360
Oy 919 gaggtatccggttgagaatgttgcttaccatgatactattcttcg 966
D 361 gaggtatccggttgagaatgttgcttaccatgatactattcttcg 408

RESULT 10
AAH50786
ID AAH50786 standard; cDNA; 401 BP.
XX
AC AAH50786;
XX


```
QY 3642 aaggagttaccttaagtgcagagaaacccaagagatgtgtgtctcttggaagcttagctt 3701
Db 181 aaggagttaccttaagtgcagagaaacccaagagatgtgtgtctcttggaagcttagctt 240
QY 3702 tctttatccaaccttaccctctcccaataagccactgtgttagttgggcccctcag 3761
Db 241 tctttatccaaccttaccctctcccaataagccactgtgttagttgggcccctcag 300
QY 3762 ggttgaagcgaagagagaagcagcagcgtttgggaaacaagactttctcctcaatagc 3821
Db 301 ggttgaagcgaagagagaagcagcagcgtttgggaaacaagactttctcctcaatagc 360
QY 3822 ctgggaaggaataaagatagatgtct 3849
Db 361 ctgggaaggaataaangataagagt 388

RESULT 12
AAH50808
ID AAH50808 standard; CDNA: 407 BP.
AC AAH50808;
XX
XX 23-AUG-2001 (first entry)
DE Human tumour associated CDNA #137.
XX
XX Human; cancer specific gene expression; gene therapy;
KM age related differential expression; ss.
XX
XX Homo sapiens.
OS
XX WO200136685-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-US31809.
XX
XX 17-NOV-1999; 99US-0166056.
XX
XX 17-NOV-1999; 99US-0166106.
XX
XX (NXYI-) NYXIS NEURO THERAPIES INC.
PA
PI Kroes RA, Moskal JR, Yamamoto H;
XX
XX WPI: 2001-355647/37.
XX
XX Novel nucleic acid molecules differentially expressed in brain cancers,
PT useful for ascertaining propensity of cell for malignant phenotype or
PT ascertaining suitability of anti-neoplastic drug candidate -
XX
XX Claim 28; Page 57; 82pp; English.
XX
XX The present invention provides the sequences of 184 CDNA fragments which
CC are differentially expressed in cancer cell depending on the age of the
CC patient. They can be used to diagnose and identify treatments for
CC cancers, particularly brain cancers such as haemangioblastoma, teratoma,
CC haemangioma, glioblastoma, schwannoma, osteoma and pinealoma. The
CC present sequence is a cancer-associated CDNA of the invention.
XX
XX Sequence 407 BP; 136 A; 73 C; 91 G; 107 T; 0 other;

Query Match 9.2%; Score 355.6; DB 22; Length 407;
Best Local Similarity 96.2%; Pred. No. 2.2e-62;
Matches 377; Conservative 0; Mismatches 9; Indels 6; Gaps 1;
```

```
Db 74 ccacttgcattcccaaatcaacaggttggccgatacccatcttgaggaacatgtctagt 133
QY 2249 tataagttccgagttggaaaagagaagcgaagctgtctaatctcatctctt 2308
Db 134 tataagttccgagttggaaaagagaagcgaagctgtctaatctcatctctt 193
QY 2309 cagtaaatattattagtaactactctgtctgttaggcattgacctgg-----gaactag 2362
Db 194 cagtaaatattattagtaactactctgtctgttaggcattgacctgggaactagaactag 253
QY 2363 agatactccaagaataacaggaaggttccctgtgctccatcaggagcttaacttccagg 2422
Db 254 agatactccaagaataacaggaaggttccctgtgctccatcaggagcttaacttccagg 313
QY 2423 gagaagagatagccaatacatagagaataatataataaggtacatgtagtgaact 2482
Db 314 gagaagagatagccaatacatagagaataatataataaggtacatgtagtgaact 373
QY 2483 gctgtggaagaaaataaagcaggggagggagt 2514
Db 374 gctgtggaagaaaataaagcaggggagggagt 405

RESULT 13
AAH1840/c
ID AAH1840 standard; CDNA: 351 BP.
AC AAH1840;
XX
XX 26-JUN-2001 (first entry)
DE Human CDNA clone (3'-primer) SEQ ID NO:8675.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length CDNA defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length CDNA -
XX
XX Claim 3; SEQ ID 8675; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length CDNA defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
```

CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SO Sequence 351 BP; 127 A; 60 C; 86 G; 67 T; 11 other;

Query Match 8.8%; Score 340; DB 22; Length 351;
 Best Local Similarity 96.9%; Pred. No. 3e-59;
 Matches 340; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1634 cccggaatttattattatagagagaggggttgcattgttgcagcagctgtgtctc 1693
 DB 351 cccggaatttattattatagagagaggggttgcattgttgcagcagctgtgtctc 292
 QY 1694 gaactctgactcagatgacatctgcttgcctccacacagctgtgattacagc 1753
 DB 291 GAACCTGTGACNMCANATGANCATNTGCTTGCCGCCACAGCTGTGGGATTACAGGC 232
 QY 1754 gttagccacacatgccgctctctctacacttaacactgtctcttactcacaact 1813
 DB 231 GTGAGCCACCATGNCGGGCTCTTTCACACTTACACCTGCTTATCCACATCT 172
 QY 1814 gtttcaacactcaatccctgctctctctcaatgttcaactctgtctccatcttag 1873
 DB 171 GTTTTACACCTTCAATCCCTGTTTCTCATGTTCACACTGTGNTCCCATGTCTATG 112
 QY 1874 ctgccttcttcaacttggtttgaaggacgtctctctgctgttcttctgtttt 1933
 DB 111 CTCGCTTTCTTACATTTGTTGTTGAAGGACAGTNTTCTGTGGCTGTGTTTGTGTTT 52
 QY 1934 cccagaaatcagatattattttaaagaagaacattcctaagaatg 1984
 DB 51 CCCAGAAATCAGTATTATTTTAAATAGAAAACATTCCTAGAAGATG 1

RESULT 14
 AAA43677/c
 ID AAA43677 standard; cDNA; 373 BP.

AC AAA43677;

DT 21-AUG-2000 (first entry)

DE Human secreted expressed sequence tag SEQ ID NO:252.

XX Human: mouse; chicken; rat; secreted expressed sequence tag; SESN;
 KM expressed sequence tag; EST; probe; chemotactic; proliferative;
 KM immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KM thrombolytic; antiinflammatory; cytosolic; antibacterial; antitumoral;
 KM antiviral; antidiabetic; antiproliferative; vulnary; antiparkinsonian;
 KM antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;
 KM cerebroprotective; anticonvulsant; antidepressant; gene therapy;
 KM vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
 KM insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KM lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KM central nervous system disorder; Alzheimer's disease; stroke;
 KM Parkinson's disease; Huntington's disease; coagulation disorder;
 KM haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
 KM tumour; infection; depression; psoriasis; ss.

OS Homo sapiens.

PN WO200021991-A1.

XX 20-APR-2000.
 PD 15-OCT-1999; 99MO-US24206.
 PF 15-OCT-1999; 98US-0104436.
 PR 15-OCT-1998; 98US-0104436.
 XX (GEM) GENETICS INST INC.
 PA JACOBS K, MCCOY JM, LAVAILLE ER, COLLINS-RACIE LA, EVANS C;
 PI Merberg D, Treacy M, Bowman MR;
 DR MPI; 2000-317938/27.
 XX Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (SESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -
 XX Claim 1; Page 267-268; 803pp; English.

CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
 CC tissue sources. The SESTs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
 CC cytosolic; antibacterial; antifungal; antiviral; antidiabetic;
 CC antistimatic; vulnary; antitumor; osteopathic; neuroprotective;
 CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene
 CC therapy and in vaccines. The SESTs are useful as probes for the
 CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.

SO Sequence 373 BP; 96 A; 121 C; 74 G; 82 T; 0 other;

Query Match 8.7%; Score 337.2; DB 21; Length 373;
 Best Local Similarity 98.9%; Pred. No. 1.e-58;
 Matches 350; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3222 ctgcagacagagatcgaatgcattgttattgttgagggtgctctcagaagaacct 3281
 DB 373 CTCGAGACAGGAGATCGAATGCACTTGTATTTGTTGGGGGCTCTCGAAGAACCT 314
 QY 3282 gttaagaccttaagcttattgttattgtctgtgaagatctctggaggtgtatcat 3341
 DB 313 GTGAAGCCCTTATACATCTTATTTAGCTGTGAAGATCTCTGGAGGTGGATCAT 254
 QY 3342 ttgaagcaagtactcaatttgaaggcaagctcttgaagaagagctgttagcatctg 3401
 DB 253 TTGAAGCAAGTACTTCAATTGAGGCAAGCTCTGTGAAGAAGGCTGAGCATCTGG 194
 QY 3402 cagctacacatcgctgttagtctgttgagggttgagggtcctggcactgctgtgtaagg 3461
 DB 193 CACCTACCAATCGCTGTGAGTGTGTGGGGCTGGGGCTCGGACACAGCTGTGTGAAGG 134
 QY 3462 gatctgacagggagaccacacagcg-ccccctacttaaacatagatgtcagtgcatitaa 3520
 DB 133 GATCTGACAGGACACACACAGCCGCCCTACTACTGAAACCATGACATGTCAGTGCATTAA 74
 QY 3521 agccatgcagctgaggggccactgagatgtctctgagattactcagagaagca 3574

Db 73 AGCCATGACGTGAGGGGCCACTGAGATGCTCTGAGATTACAAAGAATA 20

RESULT 15

AAS58007/C
ID AAS58007 standard; cDNA; 279 BP.

XX AAS58007;

DT 13-FEB-2002 (first entry)

DE cDNA #683 encoding portion of a human colon tumour protein.

XX Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.

OS Homo sapiens.

PN M0200173027-A2.

PD 04-OCT-2001.

PF 22-MAR-2001; 2001WO-US09246.

PR 24-MAR-2000; 2000US-191597P.

PR 04-MAY-2000; 2000US-202024P.

PR 05-MAY-2000; 2000US-202189P.

PA (CORI-) CORIXA CORP.

PI Meagher MJ, Xu J, King GE;

DR WPI; 2001-611627/70.

XX New colon tumour proteins and related nucleic acid, useful for

PT treatment, prevention, diagnosis and monitoring of cancer -

XX Claim 4; Page 165; 299pp; English.

XX 7th present invention relates to the isolation of novel cDNA sequences

CC encoding for at least an immunogenic portion of human colon tumour

CC proteins. The sequences of the invention are useful in pharmaceutical

CC compositions and vaccines for the prevention and treatment of cancers

CC such as colon cancer. They are also useful for the diagnosis and

CC monitoring of such cancers. Antibodies to the colon tumour proteins

CC and antigen presenting cells that express polynucleotides encoding

CC colon tumour proteins can be used to inhibit the development of

CC cancers. T-cells that react specifically with colon tumour proteins

CC are useful for removing tumour cells from samples (e.g. blood) and

CC for cancer treatment. The polynucleotide sequences are also useful in

CC gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the

CC invention that encode for portions of human colon tumour proteins.

XX Sequence 279 BP; 105 A; 45 C; 84 G; 45 T; 0 other;

Query Match 6.7%; Score 258.6; DB 23; Length 279;
Best Local Similarity 97.8%; Pred. No. 7e-43;
Matches 273; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 1681 ccaagctggtcgaactctgactcagatgacatcctgctgctccacagatgc 1740
DB 279 CCAAGCTGgtcgaactctgactcagatgacatcctgctgctccacagatgc 220
QY 1741 tggattacagagcgtgagccacatgcccgcctcttctacacttaccatgtctct 1800
DB 219 TGGATTACAGAGCGGTGAGCCACATGCCGCCCTTTCTACCTTACACTGTCTCT 160
QY 1801 tatccacacatgtcttccacact--catccctgtcttccatcgttccactgtct 1858
DB 159 TATCCTACATCTGTTTACACCTTTCATCCCTGCTCTCATGTTACACTGTCT 100
QY 1859 tcccatgttcatagtcgccttcttaccattgtgtttgaagcgacttctctgct 1918
|||||

Db 99 TCCCCATGTTTCATAGCTGCTCTTCTTACCAATTTGGTTGAAGGCGACGCTTCTGCTGCT 40

QY 1919 tgtttttgttttcccaagaatcagtatattttt 1957
|||||

Db 39 TGTTTTGTGTTTCCGAAATACGATATATTTT 1

Search completed: June 8, 2002, 00:14:17
Job time: 11567 sec

E-1.4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2002, 00:02:10 ; Search time 61.52 Seconds

(without alignments)
724.002 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130
Sequence: 1 MEGRGAAALALCVACA.....DDEGEVEYDILLEETS 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2128	99.9	401	AAV17219	Human secreted pro
2	2128	99.9	401	AAU38989	Human secreted pro
3	2121	99.6	401	AAU38989	Human secreted pro
4	572.5	26.9	804	AAU38989	Human secreted pro
5	571	26.8	736	AAU38989	Human secreted pro
6	571	26.8	736	AAU38989	Human secreted pro
7	571	26.8	736	AAU38989	Human secreted pro
8	567	26.6	736	AAU38989	Human secreted pro
9	539	25.3	363	AAU38989	Human secreted pro
10	532.5	25.0	542	AAU38989	Human secreted pro
11	532.5	25.0	747	AAU38989	Human secreted pro

SEA ID NO: 2
Database: A.Geneseq-032802
Ac. NO: AAV17219

12	530.5	24.9	359	22	AAE04200	Human gene 2 encod
13	439	20.6	592	22	AAE04244	Human gene 2 encod
14	258.5	12.1	527	22	AAU38989	Human protein sequ
15	215	10.1	60	22	AAU38989	Novel human secret
16	215	10.1	60	22	AAU38989	Novel human secret
17	199.5	9.4	105	21	AAU58500	Human polypeptide
18	198	9.3	153	22	ABU1197	Lung cancer associ
19	180	8.5	227	22	ABU1197	Human secreted pro
20	139	6.5	139	21	AAU60495	Novel human diago
21	127	6.0	173	22	AAE04217	Human secreted pro
22	100	4.7	1006	22	ABU71507	Human gene 2 encod
23	99	4.6	659	22	AAU06996	Drosophila melanog
24	96.5	4.5	404	21	AAU91465	Sequence of C3Vs.
25	96.5	4.5	489	21	AAU91609	Human secreted pro
26	96.5	4.5	544	21	AAU99373	Human secreted pro
27	96.5	4.5	544	22	AAU29160	Human PRO1374 (UNO
28	96.5	4.5	544	22	AAU29160	Human PRO polypept
29	96.5	4.5	544	22	AAU29160	Protein of the inv
30	94.5	4.4	496	16	AAU21806	Amino acid sequenc
31	94.5	4.4	816	19	AAU68094	Spliced-deleted in
32	93	4.4	1341	21	AAU48448	Mouse neuronal PAS
33	93	4.4	1358	21	AAU48448	Arabidopsis thalia
34	93	4.4	1381	21	AAU48448	Arabidopsis thalia
35	92.5	4.3	806	19	AAU75911	Arabidopsis thalia
36	92	4.3	548	14	AAU3741	Helicobacter leucy
37	91.5	4.3	309	22	ABU11152	XRL. Homo sapiens
38	91	4.3	428	20	AAU08220	Novel human diago
39	91	4.3	1173	22	AAU93171	Staphylococcus sci
40	90.5	4.2	625	22	AAU96547	Human protein sequ
41	90.5	4.2	2408	22	ABU10631	Putative P. abyssi
42	90	4.2	531	19	AAU56311	Novel human diago
43	90	4.2	629	22	ABU18854	Protein disulfide
44	89.5	4.2	561	21	AAU49740	Novel human diago
45	89.5	4.2	659	21	AAU49739	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAV17219	standard; Protein: 401 AA.
ID	AAV17219
AC	AAV17219;
DT	09-AUG-1999 (first entry)
DE	Human secreted protein (clone bd306-7).
KW	Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW	nutritional activity; cytokine; cell proliferation; immune stimulation;
KW	hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW	anti-inflammatory; tumour invasion.
OS	Homo sapiens.
FT	
FT	Misc-difference 64
FT	note="residue indicated as xaa is unspecified"
PN	WO9926961-A1.
PD	03-JUN-1999.
PF	24-NOV-1998; 98WO-US25149.
PR	23-NOV-1998; 98US-0197886.
PR	26-NOV-1997; 97US-0066804.
PA	(GENY) GENETICS INST INC.
PI	Agostino MT, Clark HF, Collins-Racie LA, Evans C;
PI	Fechtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Steininger RJ, Treacy M, Wong GG;
 XX WPI: 1999-357809/30.
 DR N-PSDB; AAX60801.
 XX
 PT New polynucleotides encoding secreted proteins
 PS Claim 10; Page 111-113; 133pp; English.
 XX
 CC The invention relates to secreted proteins (AAV17219-228) encoded by
 CC polynucleotides obtained from human fetal kidney, adult lung, adult
 CC kidney, adult brain, adult blood, adult testes, and fetal brain and
 CC murine adult bone marrow CDNA libraries. The secreted protein nucleic
 CC acid sequences (X6801-811) correspond to clones b0306-7, g1283-6,
 CC K317-3, K213-2x, na316-1, n193-20, np164-1, pe204-1, ya1-1 and yp-1,
 CC (all clones are deposited as ATCC 98599). The pns and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, hematopoiesis regulating activity,
 CC tissue growth activity, activin/ inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cachectin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The pns are
 CC also stated to be useful for gene therapy.
 XX
 SQ Sequence 401 AA;
 XX
 Query Match 99.9%; Score 2128; DB 20; Length 401;
 Best Local Similarity 100.0%; Pred. No. 4.3e-184;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEGRRGAALALALVACALRAGRAQYERYSPFRPDELPLIESAYRAHALDKYSGHW 60
 Db 1 meprgrgaalalalvacalragraqerysfrsfpdelpliesayrhaldkysghw 60
 QY 61 AESGYIEISRLRLRDLSEAFCHRNCSAARPEPAAGLASYPELRLEGGLRRACLK 120
 Db 61 aesgyieisrlrlrldseafchrnCSAARPEPAAGLASYPELRLEGGLRRACLK 120
 QY 121 RCRGGLAFRQSPSREVLADFORREPKYLOFAVYFRKANNPKALAAHTPLKHPDEM 180
 Db 121 rckrgglafRQSPSREVLADFORREPKYLOFAVYFRKANNPKALAAHTPLKHPDEM 180
 QY 181 MKRMAVYKSLPGAEDYIKDLETSYSLEFTRAVAYNGENMRTSITDMELALPDEKAF 240
 Db 181 mkrmavYKSLPGAEDYIKDLETSYSLEFTRAVAYNGENMRTSITDMELALPDEKAF 240
 QY 241 YECLAACEGSEREIDDFYISADHYVEVECKIOCEENTLPVIGYVEKEFVAATMYH 300
 Db 241 yecLAACEGSEREIDDFYISADHYVEVECKIOCEENTLPVIGYVEKEFVAATMYH 300
 QY 301 LOFAYYTLNDKNAAPCAVSYLLFDONDKVMQONTLYYYVRDHWGJSDENHFORPEAVQ 360
 Db 301 lOFAyYTLNDKNAAPCAVSYLLFDONDKVMQONTLYYYVRDHWGJSDENHFORPEAVQ 360
 QY 361 FFNVTTLQKELYDFAKENIMDDGEVEVEYVDLLELEETS 401
 Db 361 ffnvtTLQkelydfakENIMDDGEVEVEYVDLLELEETS 401
 RESULT 2
 AAU38989
 ID AAU38989 standard; Protein; 401 AA.
 AC AAU38989;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human secreted protein b0306_7.

XX
 KW Human; secreted protein; antiinflammatory; immunosuppressive;
 KW nootropic; neuroprotective; antiarthritic; antimicrobial; vulnary;
 KW cyostatic; antidiabetic; virucide; antifertility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antileuc; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW autoimmune disorder; severe combined immunodeficiency; SCID; tumour;
 KW graft-versus-host disease; multiple sclerosis; rheumatoid arthritis;
 KW periodontal disease; osteoporosis; wounding healing; ulcer;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement; vaccine.
 KW
 OS Homo sapiens.
 XX
 PN WO200175068-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 22-MAR-2001; 2001WO-US09369.
 XX
 PR 30-MAR-2000; 2000US-0539330.
 PR 04-DEC-2000; 2000US-0729674.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie E, Collins-racle LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 PI Clark H, Fechtel K, Merberg D;
 XX
 DR WPI: 2001-639363/73.
 XX
 PT N-PSDB; AAS59207.
 PT
 PT Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke) -
 XX
 PS Claim 2; Page 457-459; 619pp; English.
 XX
 CC The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
 CC in the treatment of burns, incisions and ulcers; as well as in treatment
 CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
 CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 CC infarction of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activin- or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibins to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC proteins and nucleic acids are also useful as food supplements. The
 CC present sequence represents a secreted protein of the invention.
 CC
 XX
 SQ Sequence 401 AA;
 XX
 Query Match 99.9%; Score 2128; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 4.3e-184;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPRRGAAALLALLCVCALRAGRAOYERYSFSPFDELMPLESAYRHALLDKYSGEHW 60
DB 1 meprirgaaaallallcvacalragraqerysfrsfpdelmpleasayrhaldkysgehw 60
QY 61 AESXGYLEISLRHLRLRDSFAFCHRNCSAPPEPAAGLASYPELRLFGGLRRACLK 120
DB 61 aesxygleislrhlrlrdsfaichrnscsaapqepaaglasypelrlfgglrranclk 120
QY 121 RCKGGLPAFROSQPSREVLADFQRRPEPKLQFAVYFKANNLPKATIAAHFLLKHPDEM 180
DB 121 rckgglpafqsgpsrevladfqrrepykllqfayfkannlpkaaahtllkhpddem 180
QY 181 MKRMAVYKSLPGAEDYIKOLETSYSLFIRAVRAYNGENWRSTIDMELALDPFKAF 240
DB 181 mkrmavyskslpgaedylkoletsyselfiravrayngenwrstidmelaalpdkaf 240
QY 241 YECIAACGSRKIDFDFYLSIDHVEVELECKIQCEENLTPYIGYPAVEKFAVATMYH 300
DB 241 yeciaacgsrklfdfyisidhvevelecki qceenltpyigypavekfavatyhy 300
QY 301 LQFAVYKRLNDLKNAAPCAVSYLFDQNDKVMQONLVYQYHRDTWGLSDEHFQRPPEAVQ 360
DB 301 lqfayykrlnldknaapcavsyllfdqndkvmqonlvyyqyhrdtwglstdenhfgprpeavq 360
QY 361 FFNVTTLQKELYDPAKENIMDDGEEVVEYVDLLELEETS 401
DB 361 ffnvtllqkelydfakenimddgeveveyvddlleleets 401

RESULT 3

AAB93152 standard; Protein; 401 AA.

AAB93152:

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:12067.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length cDNAs -

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

SO Sequence 401 AA;

Query Match 99.6%; Score 2121; DB 22; Length 401;

Best Local Similarity 99.5%; Pred. No. 1.8e-183;

Matches 399; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 1 MEPRRGAAALLALLCVCALRAGRAOYERYSFSPFDELMPLESAYRHALLDKYSGEHW 60
DB 1 meprirgaaaallallcvacalragraqerysfrsfpdelmpleasayrhaldkysgehw 60
QY 61 AESXGYLEISLRHLRLRDSFAFCHRNCSAPPEPAAGLASYPELRLFGGLRRACLK 120
DB 61 aesxygleislrhlrlrdsfaichrnscsaapqepaaglasypelrlfgglrranclk 120
QY 121 RCKGGLPAFROSQPSREVLADFQRRPEPKLQFAVYFKANNLPKATIAAHFLLKHPDEM 180
DB 121 rckgglpafqsgpsrevladfqrrepykllqfayfkannlpkaaahtllkhpddem 180
QY 181 MKRMAVYKSLPGAEDYIKOLETSYSLFIRAVRAYNGENWRSTIDMELALDPFKAF 240
DB 181 mkrmavyskslpgaedylkoletsyselfiravrayngenwrstidmelaalpdkaf 240
QY 241 YECIAACGSRKIDFDFYLSIDHVEVELECKIQCEENLTPYIGYPAVEKFAVATMYH 300
DB 241 yeciaacgsrklfdfyisidhvevelecki qceenltpyigypavekfavatyhy 300
QY 301 LQFAVYKRLNDLKNAAPCAVSYLFDQNDKVMQONLVYQYHRDTWGLSDEHFQRPPEAVQ 360
DB 301 lqfayykrlnldknaapcavsyllfdqndkvmqonlvyyqyhrdtwglstdenhfgprpeavq 360
QY 361 FFNVTTLQKELYDPAKENIMDDGEEVVEYVDLLELEETS 401
DB 361 ffnvtllqkelydfakenimddgeveveyvddlleleets 401

RESULT 4

AAB94482 standard; Protein; 804 AA.

AAB94482:

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:15160.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

```

PF 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 15160; 2537bp + CD ROM; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
XX Sequence 804 AA;
SQ

```

```

Db 294 kpfedflpshnylqfaynigycgavcacaktyllffpndevmqnlayeam-----1 348
QY 348 SDEHQ---PREAYOFNVTLOKELYDFAE---NIMDDDE---GPVV 388
Db 349 geehtsigrpresakeygrsilekellffaydvrfqfpdpdswtprgevi 399
XX
XX RESULT 5
XX AAB93142
XX ID AAB93142 standard: Protein; 736 AA.
XX
XX AAB93142;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12045.
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
XX Claim 8; SEQ ID 12045; 2537bp + CD ROM; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
XX Sequence 736 AA;
SQ

```


XX DT 23-MAY-2001 (first entry)
 XX XX Human membrane or secretory protein clone PSEC0109.
 DE Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes.
 XX OS Homo sapiens.
 XX PN EPI067182-A2.
 XX PD 10-JAN-2001.
 XX PE 07-JUL-2000; 2000EP-0114090.
 XX PR 08-JUL-1999; 99JP-0194179.
 XX PR 11-JAN-2000; 2000JP-0118775.
 XX PR 02-MAY-2000; 2000JP-0183766.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 DR WPI: 2001-093989/11.
 XX N-PSDB: AAF93800.
 PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX Claim 1; SEQ ID 114; 609pp + CD ROM; English.
 PS
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
 CC (ELISA)). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.
 CC
 SQ Sequence. 736 AA;

Query Match 26.8%; Score 571; DB 22; Length 736;
 Best Local Similarity 34.4%; Pred. No. 8,1e-43;
 Matches 145; Conservative 66; Mismatches 168; Indels 40; Gaps 11;

QY 6 RGAATLALLCVACALRAGRAQYERYSFRSPRDELMPLESAYRHALDKYSGEHMAESXG 65
 Db 4 ralklltlllavvaa--asqgevesaagwgmvtpdll-----faegtaayrgdwgpyvl 56
 QY 66 YLEISLRLHLRLRDSAFCHNCSA-----APOPEPAAGLASYPELIRLFGGLIRR 115
 Db 57 smeraltsraalralrlrlrcrtqcaadfpwelqpdwspspagaasaaalrldsfggllrr 116
 QY 116 AHCLKRKQGLPAPRQSGPSREVLADFGRRRPFYFLQFAHYKANNLRKAIAAHTFLK 175

Db 117 aacilrtc-lgppaahsl--seemelefrktrspnylqvaayfikinklekavaaahltfygn 173
 QY 176 PDDENMKRNMAAYKSLPGCAEDY-IRKLETKSYESLFIKRAVRAYNGENMFTSTIDNELAP 234
 Db 174 pehmengnldydyqlumsgykeadffkldetqpmgetrlgvrilyseeqpgaeayphleaalq 233
 QY 235 DEFFKAFYECLAAEGSREIKDFK-----DFYLSADHAYVEVELECKIOCEENL-TPVIGG 287
 Db 234 eyfvaeeccralcepgydygynyleynadlfqaltldhylylncqncvteleashpsre 293
 QY 288 YPEKEFVATMYHYLOFAYYKLNLDKNAPCANSTYLLFPDQNKVMOONLVYYGHHDTWGL 347
 Db 294 kpfedflpslhyngylqfayynlgnylqawecacktyllffpndevmngulayyaam-----I 348
 QY 348 SDEHFQ---PREAVQFVNTLQKELYDFAKE---NIMDDGECEVVEYVDLLELEET 400
 Db 349 geehtcrsigrpesakeygrsilekellffaydvrglfpvdpsvtrpeevlprkrlqexqk 408
 QY 401 S 401
 Db 409 s 409
 RESULT 8
 AAB36392 ID AAB36392 standard; Protein; 736 AA.
 AC AAB36392;
 DT 27-FEB-2001 (first entry)
 XX
 XX Human tumour suppressor GroS1-S protein SEQ ID NO:4.
 DE
 DE Tumour suppressor; GroS1-L; GroS1-S; cell proliferation; regulation;
 KW cancer; cytostatic; gene therapy.
 KW
 OS Homo sapiens.
 XX
 XX W0200065047-A1.
 PN
 PD 02-NOV-2000.
 XX
 XX 26-APR-2000; 2000WO-JP02731.
 PF
 XX 26-APR-1999; 99JP-0118806.
 PR
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 XX Wadhwa R, Sugihara T, Yoshida A;
 PI
 DR WPI: 2000-687340/67.
 DR N-PSDB: AAC64725.
 XX
 PS Full-length tumor suppressor genes encoding GroS1-L, GroS1-S and mouse
 PT homologs participating in regulation of cell proliferation, useful in
 PT development of preventives and remedies of cancer -
 XX
 PS Claim 1; Page 78-83; 114pp; Japanese.
 XX
 CC The present sequence represents the human tumour suppressor designated
 CC GroS1-S. GroS1-L and GroS1-S have cytostatic activity and can be used
 CC in gene therapy. GroS1-L and GroS1-S genes are useful in the development
 CC of drugs used to treat and prevent cancer.
 CC
 SQ Sequence 736 AA;

Query Match 26.6%; Score 567; DB 21; Length 736;
 Best Local Similarity 34.2%; Pred. No. 1.9e-42;
 Matches 144; Conservative 68; Mismatches 169; Indels 40; Gaps 11;

QY 6 RGAATLALLCVACALRAGRAQYERYSFRSPRDELMPLESAYRHALDKYSGEHMAESXG 65

```
Db      4 ralklltlllavvaa--asgaeveseagwgmvtprdl-----faegtaayargdwpgvvl 56
QY      66 YLEISLRHLRLDSEAFCHRNCSA-----APQPPAAGLASYPELRFGGLRR 115
Db      57 smeralrstralarlrctgcgaadfpwldpdpwspspgaagagaldrlsfifg9lllr 116
QY      116 AHLCKRCQGLPAPROSQSPREVLADQREPRKFLQFAFKANNLPKATAAHTFLKH 175
Db      117 aacrlrrc-lyppaahs--seemelefrksypynglvayfklnklekavaaahltffvgn 173
QY      176 PDDEMKRNNAVYKSLPGAEDY-IKDLETKSYESLFRAYRAYNGENWRTSITDMELAP 234
Db      174 pehmengqndyyqtsygvkeadfkldetqphmgfefflyrlyseegpqaavphleaaq 233
QY      235 DFFKAFYECLACEGSREIKDFK-----DFYLSIADHYVEVECKIOCEENT-TPYIGG 287
Db      234 eyfvaeyeccralcegydygynyleynadlfqalcldhyqvlncqncvtelashpsre 293
QY      288 YPVEKTVATMYHLYQFAFYKLNLDKNAAPCAVSYLEFDONDKYMOONLVYYOYHRDTWGL 347
Db      294 kpfedflpsnyhlyqfayynignytqagacakyllffpndevmngnlayaam-----l 348
QY      348 SDEHFO--PRPEAVQFFNVYTLQKELYDFAKE---NIMDDDEGEVEVEYVDLLELEET 400
Db      349 geehtsrsgpresakeyrtsllekellffaydvfgipfvdpdswpreevlpkrlqekq 408
QY      401 S 401
Db      409 S 409

RESULT 9
AAB36391
ID AAB36391 standard; Protein; 363 AA.
XX
AC AAB36391;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tumour suppressor Gros1-L protein SEQ ID NO:2.
XX
KW Tumour suppressor; Gros1-L; Gros1-S; cell proliferation; regulation;
KW cancer; cytosolic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200065047-A1.
XX
PD 02-NOV-2000.
XX
PF 26-APR-2000; 2000MO-JP02731.
XX
PR 26-APR-1999; 99JP-0118806.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Wadhwa R, Sugihara T, Yoshida A;
XX
DR WPI: 2000-687340/67.
DR N-PSDB: AAC64724.
XX
PT Full-length tumor suppressor genes encoding Gros1-L, Gros1-S and mouse
PT homologs participating in regulation of cell proliferation, useful in
PT development of preventives and remedies of cancer
XX
XX Claim 1; Page 67-69; 114pp; Japanese.
XX
PS The present sequence represents the human tumour suppressor designated
CC Gros1-L. Gros1-L and Gros1-S have cytosolic activity and can be used
CC in gene therapy. Gros1-L and Gros1-S genes are useful in the development
CC of drugs used to treat and prevent cancer.
XX
```

```
SO Sequence 363 AA;
Query Match 25.3%; Score 539; DB 21; Length 363;
Best Local Similarity 35.4%; Pred. No. 2,4e-40;
Matches 132; Conservative 59; Mismatches 146; Indels 36; Gaps 10;
QY      6 RGAALLALICVACALRAGRAOYERSFRSPDELMPLSAVRHLDKXSGEHMAESXG 65
Db      4 ralklltlllavvaa--asgaeveseagwgmvtprdl-----faegtaayargdwpgvvl 56
QY      66 YLEISLRHLRLDSEAFCHRNCSA-----APQPPAAGLASYPELRFGGLRR 115
Db      57 smeralrstralarlrctgcgaadfpwldpdpwspspgaagagaldrlsfifg9lllr 116
QY      116 AHLCKRCQGLPAPROSQSPREVLADQREPRKFLQFAFKANNLPKATAAHTFLKH 175
Db      117 aacrlrrc-lyppaahs--seemelefrksypynglvayfklnklekavaaahltffvgn 173
QY      176 PDDEMKRNNAVYKSLPGAEDY-IKDLETKSYESLFRAYRAYNGENWRTSITDMELAP 234
Db      174 pehmengqndyyqtsygvkeadfkldetqphmgfefflyrlyseegpqaavphleaaq 233
QY      235 DFFKAFYECLACEGSREIKDFK-----DFYLSIADHYVEVECKIOCEENT-TPYIGG 287
Db      234 eyfvaeyeccralcegydygynyleynadlfqalcldhyqvlncqncvtelashpsre 293
QY      288 YPVEKTVATMYHLYQFAFYKLNLDKNAAPCAVSYLEFDONDKYMOONLVYYOYHRDTWGL 347
Db      294 kpfedflpsnyhlyqfayynignytqagacakyllffpndevmngnlayaam-----l 348
QY      348 SDEHFO--PRPE 357
Db      349 geehtsrsgpreq 361
```

```
RESULT 10
AAB36394
ID AAB36394 standard; Protein; 542 AA.
XX
AC AAB36394;
XX
DT 27-FEB-2001 (first entry)
XX
DE Mouse tumour suppressor Gros1-S protein SEQ ID NO:8.
XX
KW Tumour suppressor; Gros1-L; Gros1-S; cell proliferation; regulation;
KW cancer; cytosolic; gene therapy.
XX
OS Mus musculus.
XX
PN WO200065047-A1.
XX
PD 02-NOV-2000.
XX
PF 26-APR-2000; 2000MO-JP02731.
XX
PR 26-APR-1999; 99JP-0118806.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Wadhwa R, Sugihara T, Yoshida A;
XX
DR WPI: 2000-687340/67.
DR N-PSDB: AAC64727.
XX
PT Full-length tumor suppressor genes encoding Gros1-L, Gros1-S and mouse
PT homologs participating in regulation of cell proliferation, useful in
PT development of preventives and remedies of cancer
XX
XX Claim 1; Page 104-107; 114pp; Japanese.
XX
PS The present sequence represents the mouse tumour suppressor designated
CC
```

CC Gros1-S, Gros1-L and Gros1-S have cytoskeletal activity and can be used
 CC in gene therapy. Gros1-L and Gros1-S genes are useful in the development
 CC of drugs used to treat and prevent cancer.
 XX
 XX Sequence 542 AA:

Query Match 25.0%; Score 532.5; DB 21; Length 542;
 Best Local Similarity 35.4%; Pred. No. 1.6e-39;
 Matches 135; Conservative 57; Mismatches 154; Indels 35; Gaps 11;

QY 48 YRHADKXSGEHWASXGYLEISRLHLRLDRSEAFCHRNCSA----APQ----PEPA-- 97
 Db 39 yaegtaaystrtdwpgvnlmeraltstraalrlrctrcatelpwabdldlgsdpals 98
 QY 98 --AGLASYPELRLEFGGLRRHCLKRCQGLPAFRQSQPSREVLADFORREPKELQPAY 155
 Db 99 qdpgaaahldlrffgavlrtracltrc-1gppsahl--seeldlefnkspynlyqay 155
 QY 156 FKANNLPKALAAHTFLKHPDDEMKRNMAVYKSLPGAEDY-1KDETKSYESLPIRAY 214
 Db 156 fkinlekavaaahltffvgnpnhmemrgnldyqqlmsgvskeadfrdleakphmhefrlqv 215
 QY 215 RAYNGENRSTIDMELALDPFKAFYECCLACSGSREIKDPK-----DYLSTADHYV 268
 Db 216 rlyseekpqeavphleaalgelyfvaecraclcegydydgnlyldysadlfgaltidhyv 275
 QY 269 EVLECKIOCEENL-TPVIGGYRVEKFTVATMYHYLOFAYYKINDLKNAAPCAVSYLLFDON 327
 Db 276 qvlnckqvcvelashpsrekrfedflpsghnylgfayngnyqalecacktyllffpn 335
 QY 328 DKVMQONLVYQYHNDTWGLSDH---FQRPPEAVQFPNVTTLQEL---YDPAKENIM 380
 Db 336 devmhqnlaly----tamlgееааsisprеаееyrrpnllekellffaydlfglprfv 390
 QY 381 DDDEGEVVEYVDLLEETS 401
 Db 391 dpdswtpeevlprkrlqekqks 411

RESULT 11
 AAB36393
 ID AAB36393 standard; Protein: 747 AA.

XX AAB36393;
 XX
 DT 27-FEB-2001 (first entry)
 XX
 DE Mouse tumour suppressor Gros1-L protein SEQ ID NO:6.
 XX
 KW Tumour suppressor; Gros1-L; Gros1-S; cell proliferation; regulation;
 KW cancer; cytoskeletal; gene therapy.
 XX
 OS Mus musculus.
 XX
 PN NO200065047-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 26-APR-2000; 2000MO-JF02721.
 XX
 PR 26-APR-1999; 99JP-0118806.
 XX
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 PI Wadhwa R, Sugihara T, Yoshida A;
 XX
 DR WPI: 2000-687340/67.
 DR N-PSDB; AAC64726.
 XX
 PT Full-length tumor suppressor genes encoding Gros1-L, Gros1-S and mouse
 PT homologs participating in regulation of cell proliferation, useful in
 PT development of preventives and remedies of cancer

XX
 PS Claim 1; Page 91-97; 114pp; Japanese.
 XX

CC The present sequence represents the mouse tumour suppressor designated
 CC Gros1-L, Gros1-L and Gros1-S have cytoskeletal activity and can be used
 CC in gene therapy. Gros1-L and Gros1-S genes are useful in the development
 CC of drugs used to treat and prevent cancer.
 XX
 XX Sequence 747 AA:

Query Match 25.0%; Score 532.5; DB 21; Length 747;
 Best Local Similarity 35.4%; Pred. No. 2.5e-39;
 Matches 135; Conservative 57; Mismatches 154; Indels 35; Gaps 11;

QY 48 YRHADKXSGEHWASXGYLEISRLHLRLDRSEAFCHRNCSA----APQ----PEPA-- 97
 Db 39 yaegtaaystrtdwpgvnlmeraltstraalrlrctrcatelpwabdldlgsdpals 98
 QY 98 --AGLASYPELRLEFGGLRRHCLKRCQGLPAFRQSQPSREVLADFORREPKELQPAY 155
 Db 99 qdpgaaahldlrffgavlrtracltrc-1gppsahl--seeldlefnkspynlyqay 155
 QY 156 FKANNLPKALAAHTFLKHPDDEMKRNMAVYKSLPGAEDY-1KDETKSYESLPIRAY 214
 Db 156 fkinlekavaaahltffvgnpnhmemrgnldyqqlmsgvskeadfrdleakphmhefrlqv 215
 QY 215 RAYNGENRSTIDMELALDPFKAFYECCLACSGSREIKDPK-----DYLSTADHYV 268
 Db 216 rlyseekpqeavphleaalgelyfvaecraclcegydydgnlyldysadlfgaltidhyv 275
 QY 269 EVLECKIOCEENL-TPVIGGYRVEKFTVATMYHYLOFAYYKINDLKNAAPCAVSYLLFDON 327
 Db 276 qvlnckqvcvelashpsrekrfedflpsghnylgfayngnyqalecacktyllffpn 335
 QY 328 DKVMQONLVYQYHNDTWGLSDH---FQRPPEAVQFPNVTTLQEL---YDPAKENIM 380
 Db 336 devmhqnlaly----tamlgееааsisprеаееyrrpnllekellffaydlfglprfv 390
 QY 381 DDDEGEVVEYVDLLEETS 401
 Db 391 dpdswtpeevlprkrlqekqks 411

RESULT 12
 AAE04200
 ID AAE04200 standard; Protein: 359 AA.

XX AAE04200;
 XX
 DT 09-AUG-2001 (first entry)
 XX
 DE Human gene 2 encoded secreted protein HUVFY29, SEQ ID NO:54.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotensin disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; gene therapy;
 KW endocrine disorder; infection; wound healing; vulnerability;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1..18
 FT Protein /label- signal_peptide
 FT 19..359 /note- "Mature secreted protein"

XX WO200179449-A2.
 PN
 XX 25-OCT-2001.
 PD
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 XX
 PS Claim 20; Page 686; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 60 AA;

Query Match 10.1%; Score 215; DB 22; Length 60;
 Best Local Similarity 73.2%; Pred. No. 4e-12;
 Matches 41; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 QY 272 ECKIQCEENTPVIIGYVKEKFAVMYHYQFAYYKLNLIKNAAPCAVSYLLFDQN 327
 DB 5 eckihckkkispgrsvenfvdmydykqpaykkindlnadpcavryllfdqn 60

Search completed: June 8, 2002, 02:08:03
 Job time: 7553 sec

THIS PAGE BLANK (15 PTO)

p. 1, 2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2002, 20:59:40 ; Search time 126.38 Seconds
(without alignments)
7523.717 Million cell updates/sec

Title: US-09-729-674-1

Perfect score: 3871

Sequence: 1 ttctctctctctctctctt.....aataaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCrUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	466.4	12.0	618	4	US-09-328-111-215 Sequence 215, App
2	228.8	5.9	9365	4	US-09-608-285A-8 Sequence 8, Appl
3	228.8	5.9	14747	4	US-09-608-285A-42 Sequence 42, Appl
4	228.8	5.9	15977	4	US-09-608-285A-59 Sequence 59, Appl
5	227	5.9	11725	2	US-08-756-506-1 Sequence 1, Appl
6	223.6	5.8	4773	3	US-08-884-324-9 Sequence 9, Appl
7	223.6	5.8	11464	3	US-08-884-324-13 Sequence 13, Appl
8	223.6	5.8	28994	3	US-08-884-324-14 Sequence 14, Appl
9	223	5.8	22481	4	US-08-367-841A-43 Sequence 43, Appl
10	223	5.8	22481	5	PCR-US95-07201-43 Sequence 43, Appl
C 11	222.8	5.7	282	1	US-08-133-629-8 Sequence 8, Appl
12	221.8	5.7	15297	4	US-09-817-180-3 Sequence 3, Appl
13	220.2	5.7	56516	2	US-08-996-306-1 Sequence 1, Appl
14	220.2	5.7	56516	4	US-09-338-907-1 Sequence 1, Appl
15	220.2	5.7	56520	4	US-09-218-207-1 Sequence 1, Appl
16	220.2	5.7	56520	4	US-09-338-907-179 Sequence 179, App
17	220.2	5.7	56520	4	US-09-218-207-179 Sequence 179, App
C 18	219.8	5.7	35060	3	US-08-814-095-7 Sequence 7, Appl
C 19	218.4	5.6	84495	4	US-09-797-906-3 Sequence 3, Appl
C 20	218	5.6	4285	4	US-09-040-774-1 Sequence 1, Appl
21	217.6	5.6	1901	4	US-09-338-907-181 Sequence 181, App
22	217.6	5.6	1901	4	US-09-218-207-181 Sequence 181, App
C 23	217.2	5.6	2061	2	US-08-960-023-11 Sequence 11, Appl
24	217.2	5.6	8453	4	US-08-167-681-45 Sequence 45, Appl
25	217.2	5.6	31571	1	US-08-323-443B-1 Sequence 1, Appl
26	217.2	5.6	53526	3	US-08-658-136-2 Sequence 2, Appl
27	217.2	5.6	53577	3	US-08-658-136-1 Sequence 1, Appl

28	216.8	5.6	72604	4	US-09-268-992-7 Sequence 7, Appl
29	216.6	5.6	35100	1	US-08-306-691B-19 Sequence 19, Appl
C 30	216.6	5.6	35100	5	PCR-US93-06251-19 Sequence 19, Appl
C 31	216.2	5.6	631	4	US-09-385-982-354 Sequence 354, App
C 32	215.8	5.6	87350	3	US-08-781-891-79 Sequence 6, Appl
C 33	215.4	5.6	2713	2	US-08-916-901-6 Sequence 6, Appl
C 34	215.4	5.6	2713	4	US-09-154-602-6 Sequence 6, Appl
35	215.4	5.6	8174	1	US-07-914-281-5 Sequence 5, Appl
36	215.4	5.6	8174	1	US-08-393-246-5 Sequence 5, Appl
37	215.4	5.6	8174	1	US-08-525-058A-5 Sequence 5, Appl
38	215.4	5.6	8174	2	US-08-696-731-5 Sequence 5, Appl
39	215.4	5.6	8174	4	US-09-042-531-5 Sequence 3, Appl
40	215.4	5.6	8174	5	PCR-US91-00899-3 Sequence 3, Appl
41	215.4	5.6	72928	3	US-09-009-913-1 Sequence 1, Appl
42	214.8	5.5	619	4	US-09-385-982-358 Sequence 358, App
C 43	214.6	5.5	2688	2	US-08-909-965C-1 Sequence 1, Appl
44	214.4	5.5	7720	4	US-09-318-448-5 Sequence 5, Appl
45	214.2	5.5	13158	2	US-08-687-080-105 Sequence 105, App

ALIGNMENTS

```
RESULT 1
US-09-328-111-215/c
; Sequence 215, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carliio, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(618)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-215

Query Match      12.0%; Score 466.4; DB 4; Length 618;
Best Local Similarity 94.8%; Pred. No. 1.le-99;
Matches 506; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

QY 1076 ccagtaacacagagacacttgggacctcgcgtgacacacacagacacagcgaagc 1135
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 534 CNAAGTCCCNAGGNCANTGGGCGCTTNGATGACGANTTCCACCCAGNCTTAANC 475
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1136 agttcagttcttaagtacac-cactccagaagagc-tgtatgacttt-gctaagga 1192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 474 AGTTCAgTCTTAAGTCAACGACCTCCAGANGAGCTTGATGACCTTTGGNTAAGA 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1193 aatatatgatgatgatgaggaagttgtgatatatgtgatgacacttgggaact 1252
```

```
Db 414 AAAATATATGATGATGATGAGGAGGAGTGTGGAATATGTGATGACCTCTTGAACT 355
Qy 1253 ggaagagaccagtagaccacagcaagcaaacagagcctcctcttgaggcttaaggaacac 1312
Db 354 GGAGAGACACAGTATGACCCACAGCACAAAGACACTTCTTGAGGCTTACGAGAAACAC 295
Qy 1313 agatctcttgctcttcccaacagccagcctgttgatacctaagagcctctcttac 1372
Db 294 AGATCTCTTGTCTTCTTCCCAACAGCCAGGCTGTGTATACCTCAGACCTCTCTCTTAC 235
Qy 1373 tctcaagtagaagggaagcccccgtctctctaactgcatgcatcaggggttgagcctg 1432
Db 234 TCTCCAAAGTAAAGGGAAGCCCGCTCTCTAATCTCAATGCAATGCAAGGAGGAGCCTG 175
Qy 1433 ccttccatcttcaacactgcaactcatgttcaacactcttccacacttctttt 1492
Db 174 CTTTCCATCTTCAACACTGCAACCTCATGTTCACACTCTTCTCACCCTTTTCTT 115
Qy 1493 gagatgagctcgtcctctcttgcccaagctgagatgcaatgcaatgcaatgcaatg 1552
Db 114 GAGATGAGTCTGCTCTCTCTGAGGCTGAGTCAATGCAATGCAATGCAATGCAATGCAATG 55
Qy 1553 caactccgctctcttggtttcaagaatctctgctgcatcagcctccagatacc 1606
Db 54 CAACCTCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1
```

RESULT 2

US-09-608-285A-8

```
Sequence 8, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Mulero, Julio
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OR INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608, 285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583, 231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557, 800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481, 238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370, 265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350, 836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273, 447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244, 444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122, 449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118, 205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9214)
```

```
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8
```

Query Match 5.9%; Score 228.8; DB 4; Length 9365;
Best Local Similarity 75.4%; Pred. No. 1.3e-43;
Matches 261; Conservative 34; Mismatches 47; Indels 4; Gaps 2;

```
Qy 1462 tttcacaccatcttctcacccttcttcttgaagatgagctcgtc-tcttgccagg 1520
Db 7413 tattattattattattattattatttttttgagagcagctgctctgttcyyrg 7472
Qy 1521 ctgagatgcaatgcaagcttctcagctcaactgcaactcgcctcttggtttcaagcaat 1580
Db 7473 ctgagatgcaatgcaagcttctcagctcaactgcaactcgcctcttggtttcaagcaat 1580
Qy 1581 tctgctgatacgtctcccgagctcctgagatgagatgagatgagatgagatgagatgagatg 1640
Db 7533 tctgctgatacgtctcccgagctcctgagatgagatgagatgagatgagatgagatgagatg 1640
Qy 1641 a---tttgaattttagaagacgggttttccatgttgcacagctgctgttcgaac 1697
Db 7593 atttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 1752
Qy 1698 tcttgaattttagaagacgggttttccatgttgcacagctgctgttcgaac 1757
Db 7653 tcttgaattttagaagacgggttttccatgttgcacagctgctgttcgaac 1757
Qy 1758 gccacacacggcgcctcttctcacccttctcacccttctcacccttctcacccttctcaccctt 1803
Db 7713 gccacacacggcgcctcttctcacccttctcacccttctcacccttctcacccttctcaccctt 1803
```

RESULT 3

US-09-608-285A-42

```
Sequence 42, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Mulero, Julio
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OR INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608, 285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583, 231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557, 800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481, 238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370, 265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350, 836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273, 447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244, 444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122, 449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118, 205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
```

```
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42
```

```
Query Match          5.9%; Score 228.8; DB 4; Length 14747;
Best Local Similarity 75.4%; Pred. No. 1.6e-43;
Matches 261; Conservative 34; Mismatches 47; Indels 4; Gaps 2;
```

```
QY 1462 tgttcacacctatcttctcaaccttttttgagatgagctcgtc-tcttgccagg 1520
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10773 tatttattattattattattatttttttgagagctcgtctgtktccyr 10832

QY 1521 ctggagtgatgagcagctctcagctcagctcagctcagctcgtctgttgatgaacat 1580
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10833 ctggagtgatgagcagctctcagctcagctcagctcagctcgtctgttgatgaacat 10892

QY 1581 tctgctgcatcagcctcccgagctcagctcagctcagctcagctcagctcagctcagctc 1640
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10893 tctgctgcatcagcctcccgagctcagctcagctcagctcagctcagctcagctcagctc 10952

QY 1641 a---ttttagttagtagagagaggttttgcacatgttgccagctgtgtctgaac 1697
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10953 atttttttatttttagtagagagaggttttgcacatgttgccagctgtgtctgaac 11012

QY 1698 tcttgactcaagatgcatcctgcttgccttgcctccacagctgctgagattacagcgctga 1757
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11013 tctgagctcaagtgatccaccctccctcgcctcccaagtgctgrrattayaggyrta 11072

QY 1758 gccacacatgccgcctcttctcacttaacacctgtctctat 1803
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11073 gccacacacgcyracctttttgtgctgtcttttttttttttttttttttttttttttttt 11118
```

```
RESULT 4
US-09-608-285A-59
Sequence 59, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
```

```
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 15977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CD39-L4/L66 Gene Sequence
NAME/KEY: CDS
LOCATION: (245)..(461)
NAME/KEY: CDS
LOCATION: (1454)..(1533)
NAME/KEY: CDS
LOCATION: (2734)..(2877)
NAME/KEY: CDS
LOCATION: (4364)..(4439)
NAME/KEY: CDS
LOCATION: (4679)..(4714)
NAME/KEY: CDS
LOCATION: (5326)..(5414)
NAME/KEY: CDS
LOCATION: (5723)..(5802)
NAME/KEY: CDS
LOCATION: (6751)..(6812)
NAME/KEY: CDS
LOCATION: (7758)..(7859)
NAME/KEY: CDS
LOCATION: (8712)..(8852)
NAME/KEY: CDS
LOCATION: (9831)..(9887)
NAME/KEY: CDS
LOCATION: (11613)..(11728)
NAME/KEY: CDS
LOCATION: (13146)..(13691)
NAME/KEY: CDS
LOCATION: (15702)..(15839)
NAME/KEY: misc_feature
LOCATION: (14871)
OTHER INFORMATION: n = a or c or g or t
US-09-608-285A-59
```

```
Query Match          5.9%; Score 228.8; DB 4; Length 15977;
Best Local Similarity 75.4%; Pred. No. 1.7e-43;
Matches 261; Conservative 34; Mismatches 47; Indels 4; Gaps 2;
```

```
QY 1462 tgttcacacctatcttctcaaccttttttgagatgagctcgtc-tcttgccagg 1520
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12003 tatttattattattattattatttttttgagagctcgtctgtktccyr 12062

QY 1521 ctggagtgatgagcagctctcagctcagctcagctcagctcagctcagctcagctcagctc 1580
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12063 ctggagtgatgagcagctctcagctcagctcagctcagctcagctcagctcagctcagctc 12122

QY 1581 tctgctgcatcagcctcccgagctcagctcagctcagctcagctcagctcagctcagctc 1640
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12123 tctgctgcatcagcctcccgagctcagctcagctcagctcagctcagctcagctcagctc 12182

QY 1641 a---ttttagttagtagagagaggttttgcacatgttgccagctgtgtctgaac 1697
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12183 atttttttatttttagtagagagaggttttgcacatgttgccagctgtgtctgaac 12242

QY 1698 tcttgactcaagatgcatcctgcttgccttgcctccacagctgtgagattacagcgctga 1757
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12243 tctgagctcaagtgatccaccctccctcgcctcccaagtgctgrrattayaggyrta 12302

QY 1758 gccacacatgccgcctcttctcacttaacacctgtctctat 1803
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12303 gccacacacgcyracctttttgtgctgtcttttttttttttttttttttttttttttttt 12348
```

RESULT 5

US-08-756-506-1
; Sequence 1, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11725 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(3520..3630, 5093..5117, 5210..5347, 5450
; LOCATION: ..5584, 8253..8395, 9269..9386, 10516..11102)
; US-08-756-506-1

Query Match 5.9%; Score 227; DB 2; Length 11725;
Best Local Similarity 85.0%; Pred. No. 3.8e-43;
Matches 266; Conservative 0; Mismatches 45; Indels 2; Gaps 1;
QY 1469 accatatttctcaaccttttttgagatgagctcgtct--ctggccagctggag 1526
DB 6361 ACTTTCTTTTCTTTTCTTTTGTGAGATGAGTTTCACTCTTGTGTCAGGCTGGAG 6420
QY 1527 tgcattgacgtctcagcctcactgcaacctcgccctctgggttcaagaattctgct 1586
DB 6421 TGCATATACGTGATCTCAGCTCACCACACCTCCGCTCTGGATTCAGAGGATTCCT 6480
QY 1587 gcatcagctcccgagtaactgagattacaaggcatgtgcacccacgcccagctaatlttg 1646
DB 6481 GCCCAGAGCTCCGAGTAGTGGGATTACAGCATGGCCACACGCCCACTAATTTTG 6540
QY 1647 tatttttagtagaagagggttttgccaatgttgccagagctggtctcgaactctgact 1706
DB 6541 TGTTTTATAGTAGAAGAGGGGTTTCTCCGTGTGCTCAAGCTGCTTCGAACTCCTGACT 6600
QY 1707 cagatgtccatctgccttgccctccacagatgctggagattacaaggcgtagagacatg 1766
DB 6601 CAGGTGATCCACCTGCTTGGCTCTCTTAAGTGTGCTGGGATTACAGGGCGTAGCCACGCG 6660

QY 1767 ccggagctcttc 1779
DB 6661 CCAGACCTCTTTC 6673
RESULT 6
US-08-884-324-9/c
; Sequence 9, Application US/08884324
; Patent No. 6060283
; GENERAL INFORMATION:
; APPLICANT: Takamori OKURA
; APPLICANT: Kakuji TORIGOE
; APPLICANT: Masashi KURIMOTO
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
; TITLE OF INVENTION: OF INDUCING THE PRODUCTION OF INTERFERON-
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,324
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185,305/96
; FILING DATE: 27-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: OKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4773 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: placenta
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 1..4773
; IDENTIFICATION METHOD: E
; US-08-884-324-9
Query Match 5.8%; Score 223.6; DB 3; Length 4773;
Best Local Similarity 85.3%; Pred. No. 1.5e-42;
Matches 261; Conservative 0; Mismatches 44; Indels 1; Gaps 1;
QY 1476 ttltcaacttttttgagatgagctcgtctc--tttgccaggttgagtgcaatg 1554
DB 1673 TTTTCTTTTCTTTTGTGAGCCGAGTTTGTCTTGTGCTCCGCTGGAGTCAATGG 1614
QY 1535 cagcttctcagctcaacgcaacctccgctcttggtgtcaagcaatctctgcatcagc 1594
DB 1613 CAGCATATGAGCTCATGCGAACAACCTCTGCTCCGCGTTCAAAAGATTCCTGCTGACG 1554
QY 1595 ctcccgagtaactgagattacaaggcatgtgcacacacgcccagctaatlttgattta 1654

Db - 1553 CTCGCCAGTAGCTGGATTACAGCATATGCCACAGCCACACTATTTGATTTT 1494
QY 1655 gtagagacggggttttgcattgttgcacggcgtgtctcgaactcttgactagatgat 1714
Db 1493 GTAGAGATGGGTTTCTCCATGTTGGTCAGGCTGTGTGAACCTCGACTTCAGGTGAT 1434
QY 1715 ccattcgtcttgccctcccaagtgctgagattacaggcgtgagccacatgccggcct 1774
Db 1433 CCCCCCCTCGGCTCCCAAGTGTGGATTACAGGTGTGAGCCACAGCCAGCCA 1374
QY 1775 cttctc 1780
Db 1373 CATTC 1368

RESULT 7
US-08-884-324-13/c
; Sequence 13, Application US/08884324
; Patent No. 6060283
; GENERAL INFORMATION:
; APPLICANT: Takamori OKURA
; APPLICANT: Masahiro KURIMOTO
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
; OF INDUCING THE PRODUCTION OF INTERFERON-
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,324
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185,305/96
; FILING DATE: 27-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: OKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: placenta
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..3
; IDENTIFICATION METHOD: E
; NAME/KEY: leader peptide
; LOCATION: 4..82
; IDENTIFICATION METHOD: S
; NAME/KEY: intron
; LOCATION: 83..1453
; IDENTIFICATION METHOD: E

NAME/KEY: leader peptide
LOCATION: 1454..1465
IDENTIFICATION METHOD: S
NAME/KEY: intron
LOCATION: 1466..4848
IDENTIFICATION METHOD: E
NAME/KEY: leader peptide
LOCATION: 4849..4865
IDENTIFICATION METHOD: S
NAME/KEY: mat peptide
LOCATION: 4866..4983
IDENTIFICATION METHOD: S
NAME/KEY: intron
LOCATION: 4984..6317
IDENTIFICATION METHOD: E
NAME/KEY: mat peptide
LOCATION: 6318..6451
IDENTIFICATION METHOD: S
NAME/KEY: intron
LOCATION: 6452..11224
IDENTIFICATION METHOD: E
NAME/KEY: mat peptide
LOCATION: 11225..11443
IDENTIFICATION METHOD: S
NAME/KEY: 3'UTR
LOCATION: 11444..11464
IDENTIFICATION METHOD: E
US-08-884-324-13

Query Match 5.8%; Score 223.6; DB 3; Length 11464;
Best Local Similarity 85.3%; Pred. No. 2.3e-42;
Matches 261; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

QY 1476 ttctcactttttttgagatgagctcgtc-tcttgcacagcgtgagtgcaatg 1534
Db 8124 TTTTCTTTTCTTTTGTGAGACCGAGTTTGTCTTGTGCCCCAGGCTGAGTGCAATGG 8065
QY 1535 caagttcctcagctcactgaacctccgctcttggttcaagaattctgctcagc 1594
Db 8064 CACGATCTAGGCTCAGTCACCTGCTCCCGGCTCAAAAAGATTCTCTCCACG 8005
QY 1595 ctcccgatccttgagattacagcagtgccacacagcccgactaatttgatttca 1654
Db 8004 CTCGCCAGTAGCTGGATTACAGCATATGCCACAGCCACACTATTTGATTTT 7945
QY 1655 gtagagacggggttttgcattgttgcacggcgtgtctcgaactcttgactagatgat 1714
Db 7944 GTAGAGATGGGTTTCTCCATGTTGGTCAGGCTGTGTGAACCTCGACTTCAGGTGAT 7885
QY 1715 ccattcgtcttgccctcccaagtgctgagattacaggcgtgagccacatgccggcct 1774
Db 7884 CCCCCCCTCGGCTCCCAAGTGTGGATTACAGGTGTGAGCCACAGCCAGCCA 7825
QY 1775 cttctc 1780
Db 7824 CATTC 7819

RESULT 8
US-08-884-324-14/c
; Sequence 14, Application US/08884324
; Patent No. 6060283
; GENERAL INFORMATION:
; APPLICANT: Takamori OKURA
; APPLICANT: Masahiro KURIMOTO
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
; OF INDUCING THE PRODUCTION OF INTERFERON-
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300

```

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,324
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 185,305/96
FILING DATE: 27-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: OKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 28994 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: Placenta
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..15606
IDENTIFICATION METHOD: E
NAME/KEY: leader peptide
LOCATION: 15607..15685
IDENTIFICATION METHOD: S
NAME/KEY: Intron
LOCATION: 15686..17056
IDENTIFICATION METHOD: E
NAME/KEY: leader peptide
LOCATION: 17057..17068
IDENTIFICATION METHOD: S
NAME/KEY: Intron
LOCATION: 17069..20451
IDENTIFICATION METHOD: E
NAME/KEY: leader peptide
LOCATION: 20452..20468
IDENTIFICATION METHOD: S
NAME/KEY: mat peptide
LOCATION: 20469..20586
IDENTIFICATION METHOD: S
NAME/KEY: Intron
LOCATION: 20587..21920
IDENTIFICATION METHOD: E
NAME/KEY: mat peptide
LOCATION: 21921..22054
IDENTIFICATION METHOD: S
NAME/KEY: Intron
LOCATION: 22055..26827
IDENTIFICATION METHOD: E
NAME/KEY: mat peptide
LOCATION: 26828..27046
IDENTIFICATION METHOD: S
NAME/KEY: 3'UTR
LOCATION: 27047..28994
IDENTIFICATION METHOD: E
US-08-884-324-14

```

```

Query Match          5.8%; Score 223.6; DB 3; Length 28994;
Best Local Similarity 85.3%; Pred. No. 3.7e-42;
Matches 261; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

Qy 1476 ttccacccttttttttgatgagatgctc-tctgcccaagctgagtgcaatg 1534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23727 tttttttttttttttttagaccgagtttgccttcttgccacagctgagtgcaatg 23668

Qy 1535 cagctccagctcagctcagctcagctcagctcagctcagctcagctcagctcag 1594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23667 CAGCATCTAGGCTCATGCCAACCTGCTCCCGGGTTCAAAAGATTCTCCTCAGC 23668

Qy 1595 ctcccgagctacgtggaattacaggaatgtgccaccagcccgagctaatltgatttla 1654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23607 CTCGCGAGTACTGGAGTTACAGCATATGCCACACGCCAGCTAATTTGTATTTTTA 23548

Qy 1655 gtagagacggggttttgcaatgttggcagagctgtgtccgaacttctacttcagatg 1714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23547 GTAGAGATGGGGTTTCTCATGTTGTCAGGCTGGTCTTGAACCTTCACCTCAGTGAT 23488

Qy 1715 ccactgcttggcctcccaagatgctggaattacagcgtagccacacatgcccgact 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23487 CCGCCCCCTGGCCTCCCAAGTGTGCTGATTTACAGGTGTGACGACGACGCCAGCCA 23428

Qy 1775 cttctc 1780
    ||| |||
Db 23427 CATCTC 23422

```

```

RESULT 9
US-08-367-841A-43
; Sequence 43, Application US/08367841A
; Patent No. 6319687
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof.
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367, 841A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA: 07/952,796
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs

```



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/133,629
APPLICATION NUMBER: US/08/133,629
FILING DATE: 07-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Greer, Helen
REGISTRATION NUMBER: 36,816
REFERENCE/DOCKET NUMBER: M0828/7001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX: 92-1742 EZEKTEL
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-133-629-8

```

Query Match 5.8%; Score 222.8; DB 1; Length 282;

Best Local Similarity 87.6%; Pred. No. 5.3e-43;
Matches 247; Conservative 6; Mismatches 28; Indels 1; Gaps 1;

```

QY 1493 gagatggagctcgcctctctcccaagctgagtgcaatgcaagcttcacgctcaccg 1552
DB 282 GAGATGAGACTCTYRCTCTGTTGCCAGGCTGGAGTGCATGTGCGGATCWCGGCTCACG 223
QY 1553 caactccgcctctgggtgttcaagaatctctgcatcagcctcccgagtaaccggagat 1612
DB 222 CARCTCCACCTCCGCGGTCAAGGATTTCTCTGCTCAGCCTCCCGATGCTGGGAT 163
QY 1613 taacagcatgtgcacacacgcgcgctaa-tttgtatttttagtagaagcgggttttg 1671
DB 162 TACAGCGCGCGCCACACAGCGCGCTAATTTTGTATTTTGTAGTAGAGCGGGTTTCA 103
QY 1672 ccattgtggcagagctgtcttgaactcttgacttcagatgcatcctgcttgagctc 1731
DB 102 CCATGTTGGCGAGGCTGTGACTTGAACCTCTGACTGATGCCACCCACCTCGGCC 43
QY 1732 ccacatcgtgagattacagcgtgagccacatgcccggcc 1773
DB 42 CCAAGTGTGGGATTACAGTGTGAGCCACCCACCGCCAGCC 1

```

RESULT 12
US-09-817-180-3/c

```

Sequence 3, Application US/09817180
Patent No. 6340584
GENERAL INFORMATION:
APPLICANT: GAN, weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C0001183
CURRENT APPLICATION NUMBER: US/09/817,180
CURRENT FILING DATE: 2001-03-27
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 3
LENGTH: 15297
TYPE: DNA
ORGANISM: Human
US-09-817-180-3

```

Query Match 5.7%; Score 221.8; DB 4; Length 15297;

```

Best Local Similarity 81.4%; Pred. No. 7.1e-42;
Matches 281; Conservative 0; Mismatches 62; Indels 2; Gaps 2;
QY 1445 taacacctgcacacctcaatgttcaacacctctctcacccttttttttagatgagctc 1504
DB 12156 TCACCACACCCAGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 12097
QY 1505 cgcctc-tcttgcacagctgagtgcaatgcaagcttctcagctcactgcacctccgcc 1563
DB 12096 CACTCTTGTGCGCCAGGCTGAGTGACAGGACATCTCGGCTCAGCTCAGCTCCGCC 12037
QY 1564 tcttggttcaagcaattctgtctgcatcagcctcccgagtlacttggtatgaagcagct 1623
DB 12036 TCCTGGGTTCAAGCAATTTCTCCACCTCAGCTCCCAATGCTGGATTACAGGACAGT 11977
QY 1624 gccacacgcccggctaa-tttgtatttttagtagaagcgggttttgcatgttgcc 1682
DB 11976 GCCACCATGCTCGGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 11917
QY 1683 agcctggtctcgaactcttgacttaagatgatalcatalctgcttgcctccacagtgctg 1742
DB 11916 TGGCTGTCTCAAACTCCGACACAGGTGATCCACCTGCGGCTCCCAAGTGTCTG 11857
QY 1743 gattacagcgctgagccacatgcccggcccttctcacttc 1787
DB 11856 GGATTACAGGTGTGAGCCAGCTGCCCACTCTAATTTTGTATTTT 11812

```

RESULT 13
US-08-996-306-1

```

Sequence 1, Application US/08996306
Patent No. 5943522
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Chumakov, Ilya
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate cancer gene
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESS: Knobb, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,306
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET.018A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 56516 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Promoter

```

1	IDENTIFICATION:	1629..1870	PROSCAN
2	IDENTIFICATION METHOD:		
3	FEATURE:		
4	NAME/KEY:	Potential ATG	
5	LOCATION:	1998..2000	
6	FEATURE:		
7	NAME/KEY:	Exon 1	
8	LOCATION:	2001..2216	
9	FEATURE:		
10	NAME/KEY:	ATG	
11	LOCATION:	2031..2033	
12	FEATURE:		
13	NAME/KEY:	Tyr phos	
14	LOCATION:	11694..14332	
15	FEATURE:		
16	NAME/KEY:	SEQ ID42	
17	LOCATION:	11930..11947	
18	FEATURE:		
19	NAME/KEY:	SEQ ID24	
20	LOCATION:	12057..12103	
21	FEATURE:		
22	NAME/KEY:	SEQ ID51	
23	LOCATION:	comp1(12339..12358)	
24	FEATURE:		
25	NAME/KEY:	SEQ ID64	
26	LOCATION:	13547..13564	
27	FEATURE:		
28	NAME/KEY:	SEQ ID58	
29	LOCATION:	13657..13703	
30	FEATURE:		
31	NAME/KEY:	SEQ ID67	
32	LOCATION:	comp1(13962..13981)	
33	FEATURE:		
34	NAME/KEY:	Exon 2	
35	LOCATION:	18196..18265	
36	FEATURE:		
37	NAME/KEY:	Exon 3	
38	LOCATION:	23717..23832	
39	FEATURE:		
40	NAME/KEY:	Exon 4	
41	LOCATION:	25571..25660	
42	FEATURE:		
43	NAME/KEY:	SEQ ID43	
44	LOCATION:	34216..34234	
45	FEATURE:		
46	NAME/KEY:	SEQ ID25	
47	LOCATION:	34469..34515	
48	FEATURE:		
49	NAME/KEY:	SEQ ID52	
50	LOCATION:	comp1(34625..34645)	
51	FEATURE:		
52	NAME/KEY:	Exon 5	
53	LOCATION:	34659..34759	
54	FEATURE:		
55	NAME/KEY:	Exon 6	
56	LOCATION:	40688..40846	
57	FEATURE:		
58	NAME/KEY:	Exon 7	
59	LOCATION:	48070..48193	
60	FEATURE:		
61	NAME/KEY:	Exon 8	
62	LOCATION:	50182..54523	
63	FEATURE:		
64	NAME/KEY:	SEQ ID65	
65	LOCATION:	51149..51168	
66	FEATURE:		
67	NAME/KEY:	SEQ ID59	
68	LOCATION:	51448..51494	
69	FEATURE:		
70	NAME/KEY:	SEQ ID68	
71	LOCATION:	comp1(51482..51499)	
72	FEATURE:		
73	NAME/KEY:	SEQ ID44	

```

: LOCATION: 51596..51613
: FEATURE:
: NAME/KEY: SEQ ID26
: LOCATION: 51612..51658
: FEATURE:
: NAME/KEY: SEQ ID35
: LOCATION: compl(51596..52015)
: FEATURE:
: NAME/KEY: polyad signal
: LOCATION: 54445..54450
: OS-08-996-306-1

```

Query Match	5.7%;	Score 220.2;	DB 2;	Length 56516;
Best Local Similarity	76.5%;	Pred. No. 3.3e-41;		
Matches 270;	Conservative	0;	Mismatches 83;	Indels 0;

QY	1435	tttcacatctccacacgtccaaccccaagttcaaacctatcttccaccttltttg	1494
Db	20669	ttttcttccctttaaactttaaagatcttacttttgacaaacattttgttttttttga	20722
QY	1495	gattgagatctgcctctcttctcccaagctctgagtgatgaatgacgtcttcagctcaetgca	1554
Db	20729	gattgagattttcaattctttgttgccacagcctggctgttaagtgatgcataatcttcagctcacatgcga	20786
QY	1555	acctccgcctcttggtttcaagcaattctctgcgtcaagctctccagatccctggatga	1614
Db	20789	accttttgccctcccggttttaaacctatttttccgtccctcagcctcccgatagcttgggattna	20844
QY	1615	caggtcatgtgcaaccacagcccggtcaatttgtatattttagtagagacgggttttgcca	1674
Db	20849	cagacatgacacacacacgccttgcttaattttgtatttttcaatgaagatggggttttaccca	20906
QY	1675	tgattgcacagctgtgctcgtcaactttagctacatgatcatcatctgctttggccttcca	1734
Db	20909	tgttttgatcagagctgggtcttgaacctctgaacctcagatgatattgccttgccttgcgccttcccca	20966
QY	1735	cagtgctcgtgattacaagctcgtgagccacacatgcccgcgactctttctacacttt	1787
Db	20969	aactgtcagggattttacagcgttaggcacactgtgcccctctgcctctctatattttggccctttt	21021

```

RESULT 14
US-09-338-907-1
Sequence 1, Application US/09338907
Patent No. 6265546
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilyu, Chumakov
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: PROSTATE CANCER GENE
FILE REFERENCE: GENSET.18CPICP
CURRENT APPLICATION NUMBER: US/09/338,907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218,207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 56516
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: 1629..1870
OTHER INFORMATION: Identification method Proscan
FEATURE:
NAME/KEY: misc.feature

```

LOCATION: 1998..2000
OTHER INFORMATION: potential start codon
FEATURE:
NAME/KEY: exon
LOCATION: 2001..2216
OTHER INFORMATION: exon1
FEATURE:
NAME/KEY: misc-feature
LOCATION: 2031..2033
OTHER INFORMATION: ATG
FEATURE:
NAME/KEY: misc-feature
LOCATION: 11694..14332
OTHER INFORMATION: Tyr Phos
FEATURE:
NAME/KEY: primer_bind
LOCATION: 11930..11947
OTHER INFORMATION: upstream amplification primer 4-77 SEQ ID42
FEATURE:
NAME/KEY: allele
LOCATION: 12057..12103
OTHER INFORMATION: polymorphic fragment 4-77 SEQ ID24
FEATURE:
NAME/KEY: primer_bind
LOCATION: 12339..12358
OTHER INFORMATION: downstream amplification primer 4-77 SEQ ID51, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13547..13564
OTHER INFORMATION: upstream amplification primer 4-73 SEQ ID64
FEATURE:
NAME/KEY: allele
LOCATION: 13657..13703
OTHER INFORMATION: polymorphic fragment 4-73 SEQ ID58
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13962..13981
OTHER INFORMATION: downstream amplification primer 4-73 SEQ ID67, complement
FEATURE:
NAME/KEY: exon
LOCATION: 18196..18265
OTHER INFORMATION: exon 2
FEATURE:
NAME/KEY: exon
LOCATION: 23717..23832
OTHER INFORMATION: exon 3
FEATURE:
NAME/KEY: exon
LOCATION: 25571..25660
OTHER INFORMATION: exon 4
FEATURE:
NAME/KEY: primer_bind
LOCATION: 34216..34234
OTHER INFORMATION: upstream amplification primer 99-217 SEQ ID43
FEATURE:
NAME/KEY: allele
LOCATION: 34469..34515
OTHER INFORMATION: polymorphic fragment 99-217 SEQ ID25
FEATURE:
NAME/KEY: primer_bind
LOCATION: 34625..34645
OTHER INFORMATION: downstream amplification primer 99-217 SEQ ID52, complement
FEATURE:
NAME/KEY: exon
LOCATION: 34669..34759
OTHER INFORMATION: exon 5
FEATURE:
NAME/KEY: exon
LOCATION: 40688..40846
OTHER INFORMATION: exon 6
FEATURE:
NAME/KEY: exon
LOCATION: 48070..48193

OTHER INFORMATION: exon 7
FEATURE:
NAME/KEY: exon
LOCATION: 50182..54523
OTHER INFORMATION: exon 8
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51149..51168
OTHER INFORMATION: upstream amplification primer 4-65 SEQ ID65
FEATURE:
NAME/KEY: allele
LOCATION: 51448..51494
OTHER INFORMATION: polymorphic fragment 4-65 SEQ ID59
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51482..51499
OTHER INFORMATION: downstream amplification primer 4-65 SEQ ID68, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51596..51613
OTHER INFORMATION: upstream amplification primer 4-67 SEQ ID44
FEATURE:
NAME/KEY: allele
LOCATION: 51612..51658
OTHER INFORMATION: polymorphic fragment 4-67 SEQ ID26
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51996..52015
OTHER INFORMATION: downstream amplification primer 4-67 SEQ ID53, complement
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 54445..54450
OTHER INFORMATION: AATAAA
US-09-338-907-1

Query Match 5.7%; Score 220.2; DB 4; Length 56516;
Best Local Similarity 76.3%; Pred. No. 3.3e-41;
Matches 270; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1435 ttccatcttcacacgtgcacccatgcttcacacatcttcacacatttcacatttttga 1494
DB 20669 ttctcccttacccttcccttcaacgacatcttcacacacatttcttttttga 20728
QY 1495 gatggagctcgtctctcttccagcgttgagtgcaatggcagcttcacactactga 1554
DB 20729 gatggagctcgtctcttcttgcacagcgttgagtgcaatggcagcttcacactga 20788
QY 1555 acctcgctcttggtgtcaagcaattctgtcgtacatcagccctccgagtagctggatta 1614
DB 20789 accttgcctccggtgtcaagcaattctgtcgtacatcagccctccgagtagctggatta 20848
QY 1615 caggcatgtgcacacacacgctgtaatttgaatttgaatttgaaggcgggtttgcca 1674
DB 20849 cagacatgtgcacacacacacgctgtaatttgaatttgaatttgaaggcgggtttgcca 20908
QY 1675 ttttggcagcgtgtctcgaactctgactcagatgatcacaaccttgcctgcctccca 1734
DB 20909 ttttggcagcgtgtctcgaactctgactcagatgatcacaaccttgcctgcctccca 20968
QY 1735 cagtgctggattacagcgtgacacacatgcctgcctcttctcactt 1787
DB 20969 aagtgaggtattacagcgtgacacacatgcctgcctcttctcactt 21021

RESULT 15
US-09-218-207-1
; Sequence 1, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marica
; APPLICANT: Ilya, Chumakov

APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate cancer gene
FILE REFERENCE: GENSET.018Cp1
CURRENT APPLICATION NUMBER: US/09/218,207
CURRENT FILING DATE: 1998-12-22
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 56516
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: promoter
LOCATION: 1629..1870
OTHER INFORMATION: Identification method Proscan
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1998..2000
OTHER INFORMATION: potential start codon
FEATURE:
NAME/KEY: exon
LOCATION: 2001..2216
OTHER INFORMATION: exon1
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2031..2033
OTHER INFORMATION: ATG
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11694..14332
OTHER INFORMATION: Tyr Phos
FEATURE:
NAME/KEY: primer_bind
LOCATION: 11930..11947
OTHER INFORMATION: upstream amplification primer 4-77 SEQ ID42
FEATURE:
NAME/KEY: allele
LOCATION: 12057..12103
OTHER INFORMATION: polymorphic fragment 4-77 SEQ ID24
FEATURE:
NAME/KEY: primer_bind
LOCATION: 12339..12358
OTHER INFORMATION: downstream amplification primer 4-77 SEQ ID51, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13547..13564
OTHER INFORMATION: upstream amplification primer 4-73 SEQ ID64
FEATURE:
NAME/KEY: allele
LOCATION: 13657..13703
OTHER INFORMATION: polymorphic fragment 4-73 SEQ ID58
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13962..13981
OTHER INFORMATION: downstream amplification primer 4-73 SEQ ID67, complement
FEATURE:
NAME/KEY: exon
LOCATION: 18196..18265
OTHER INFORMATION: exon 2
FEATURE:
NAME/KEY: exon
LOCATION: 23717..23832
OTHER INFORMATION: exon 3
FEATURE:
NAME/KEY: exon
LOCATION: 25571..25660
OTHER INFORMATION: exon 4
FEATURE:
NAME/KEY: primer_bind

LOCATION: 34216..34234
OTHER INFORMATION: upstream amplification primer 99-217 SEQ ID43
FEATURE:
NAME/KEY: allele
LOCATION: 34469..34515
OTHER INFORMATION: polymorphic fragment 99-217 SEQ ID25
FEATURE:
NAME/KEY: primer_bind
LOCATION: 34625..34645
OTHER INFORMATION: downstream amplification primer 99-217 SEQ ID52, complement
FEATURE:
NAME/KEY: exon
LOCATION: 34669..34759
OTHER INFORMATION: exon 5
FEATURE:
NAME/KEY: exon
LOCATION: 40688..40846
OTHER INFORMATION: exon 6
FEATURE:
NAME/KEY: exon
LOCATION: 48070..48193
OTHER INFORMATION: exon 7
FEATURE:
NAME/KEY: exon
LOCATION: 50182..54523
OTHER INFORMATION: exon 8
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51149..51168
OTHER INFORMATION: upstream amplification primer 4-65 SEQ ID65
FEATURE:
NAME/KEY: allele
LOCATION: 51448..51494
OTHER INFORMATION: polymorphic fragment 4-65 SEQ ID59
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51482..51499
OTHER INFORMATION: downstream amplification primer 4-65 SEQ ID65
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51596..51613
OTHER INFORMATION: upstream amplification primer 4-67 SEQ ID44
FEATURE:
NAME/KEY: allele
LOCATION: 51612..51658
OTHER INFORMATION: polymorphic fragment 4-67 SEQ ID26
FEATURE:
NAME/KEY: primer_bind
LOCATION: 51996..52015
OTHER INFORMATION: downstream amplification primer 4-67 SEQ ID53, complement
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 54445..54450
OTHER INFORMATION: AATTAAA
US-09-218-207-1

Query Match 5.7%; Score 220.2; DB 4; Length 56516;
Best Local Similarity 76.5%; Pred. No. 3.3e-41;
Matches 270; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 1435 ttccatcttcacagctccacactatgttcaacactatcttcacattttttga 1494
DB 20669 ttcttcctccctatactcctcaagatctactttgataaacaacttgcctttttttga 20728
QY 1495 gatgagatctgcgtctcttgcacagctgagtgacatgacagcttctcagctcactga 1554
DB 20729 gatgagatcttcactctgttgcctcagctggtgtaatgttgaatcctcagctcactga 20788
QY 1555 acctccgctcttgggttcaagcaattctgctgcatcagctcccgagtaacctggagatta 1614
DB 20789 accttgcctcccggttcaagcaatttctcgtcctcagctcccgagtaagctggagatta 20848

[illegible]

```
Search completed: June 8, 2002, 00:10:10
Job time: 11430 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2002, 02:08:55 ; Search time 89.01 Seconds
(without alignments)
779.361 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130

Sequence: 1 MERGREGAALLALLCVACA.....DDEGEVEYVDLLELEETS 401

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rotent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1038	48.7	222	11 Q9DBW2	Q9DBW2 mus musculu
2	572.5	26.9	804	4 Q96SN3	Q96SN3 homo sapien
3	571	26.8	736	4 Q96SL5	Q96SL5 homo sapien
4	571	26.8	736	4 Q96SK8	Q96SK8 homo sapien
5	567	26.6	742	4 Q9HC86	Q9HC86 homo sapien
6	545	25.6	736	11 Q9CWR5	Q9CWR5 mus musculu
7	539	25.3	363	4 Q9HC87	Q9HC87 homo sapien
8	534.5	25.1	728	11 Q9RLJ8	Q9RLJ8 ratuS norv
9	532.5	25.0	742	11 Q9QZT6	Q9QZT6 mus musculu
10	532.5	25.0	747	11 Q9QZT7	Q9QZT7 mus musculu
11	372	17.5	476	5 Q9NA33	Q9NA33 caenorhabdi
12	258.5	12.1	527	4 Q9NV12	Q9NV12 homo sapien
13	161.5	7.6	545	11 Q98836	Q98836 mus musculu
14	161.5	7.6	551	4 Q15740	Q15740 homo sapien
15	159.5	7.5	551	4 Q13512	Q13512 homo sapien
16	118.5	5.6	598	5 Q9BPP2	Q9BPP2 caenorhabdi

17	104	4.9	317	16 Q97R54	Q97R54 streptococc
18	104	4.9	853	16 Q67796	Q67796 aquifex aeo
19	100.5	4.7	688	12 Q91TS4	Q91TS4 lupala herp
20	100	4.7	1006	5 Q91090	Q91090 drosophila
21	100	4.7	1006	5 Q9VUC7	Q9VUC7 drosophila
22	99	4.6	1388	6 Q9GL21	Q9GL21 canis fami
23	98.5	4.6	1575	16 Q9POH4	Q9POH4 ureaplasma
24	98	4.6	455	5 Q16626	Q16626 caenorhabdi
25	98	4.6	484	16 Q06968	Q06968 bacillus su
26	98	4.6	4202	11 Q91V63	Q91V63 mus musculu
27	98	4.6	4731	11 Q91XP8	Q91XP8 mus musculu
28	97.5	4.6	910	3 Q59796	Q59796 schizosacch
29	97.5	4.6	1162	16 Q9PMG0	Q9PMG0 campylobact
30	96	4.5	788	17 Q58603	Q58603 pyrococcus
31	96	4.5	3477	11 Q9NMH8	Q9NMH8 mus musculu
32	96	4.5	4202	11 Q91XP9	Q91XP9 mus musculu
33	96	4.5	4731	11 Q91X00	Q91X00 mus musculu
34	95.5	4.5	240	17 Q981E0	Q981E0 sulfolobus
35	95.5	4.5	336	5 Q95OS2	Q95OS2 caenorhabdi
36	94	4.4	435	16 Q9PMV0	Q9PMV0 campylobact
37	94	4.4	439	5 Q9GSD0	Q9GSD0 plasmodium
38	93.5	4.4	541	5 Q95P11	Q95P11 brugia mala
39	93	4.4	434	16 Q97KG2	Q97KG2 clostridium
40	93	4.4	757	10 Q9F174	Q9F174 arabidopsis
41	92.5	4.3	729	16 Q9PHP5	Q9PHP5 campylobact
42	92	4.3	411	16 Q9HV83	Q9HV83 pseudomonas
43	92	4.3	534	3 Q9USM8	Q9USM8 schizosacch
44	92	4.3	877	2 Q9RG09	Q9RG09 neisseria m
45	91.5	4.3	429	2 Q91AL2	Q91AL2 moraxella c

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	222 AA.
Q9DBW2	Q9DBW2	Q9DBW2		
AC	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	5730529N23RIK PROTEIN.			
GN	CRTP OR 5730529N23RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochias H.,			
RA	Kuethl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombaris P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL, AK007624; BAB25144.1; -			
DR	MGD; MGI:1891221; Crtrap.			

SQ SEQUENCE 222 AA; 25803 MW; 336DF2019A126B04 CRC64;

Query Match	48.7%	Score 1038	DB 11	Length 222
Best Local Similarity	86.5%	Pred. No. 2.4e81		
Matches 192	Conservative 13	Mismatches 17	Indels 0	Gaps 0

QY	180	MMNNMAYKSLPCAGEDYIKOLEKKSYESLPIRAVVRVNGNMWSTSTDMELALPDEFKA	23
		: : : : : :	
		1 MMRKRRYKSLPCAGABDIKOLEKKSYESLPIRAVVRVNGNMWSTSTDMELPLPFLKA	60
QY	240	FYECLAAEGSREIKDFEFLSTADHYVEVLECKIOCEENLPIVIGYEVKEKVAAMYH	269
		: : : : : :	
Db	61	FYECLAAEGVAGNOGILQDFFLSTADHYVGSLECKINCEBTLPIVIGYEVKEKVPIMYH	120
QY	300	YLOFAYYKLNLDKNAAPCAVSYLLFDQNDKVMQONLVIYQYHRDTWGLSDENHQPREAV	359
		: : : : : :	
Db	121	YLOFAYYKLNLDKNAAPCAVSYLLFDQSDRVMQONLVIYQYHRDKWGLSDENHQPREAV	180
QY	360	QEFNVTTLOKELYDPAKENIMDDDEGEVEVYVDLLELEFS	401
		: : : : : :	
Db	181	QEFNVTTLOKELYDPAQENIMDDDEGEVEVYVDLLELEESA	222

RESULT	2			
ID	0965N3	PRELIMINARY:	PRT:	804 AA.
AC	0965N3;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, last annotation update)			
DE	CNNA FL114742 FIS. CLONE NT2RP3002650, HIGHLY SIMILAR TO MUS			
DE	MUSCULUS GROWTH SUPPRESSOR 1L (GROSL) mRNA.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Isoegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi			
RA	Nakanabe J., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.			
RA	Yamamoto J., Wakamatsu K., Nakamura Y., Nagahari K., Masuno Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	"MEDO human cDNA sequencing project."			
RL	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
RL	EMBL: AK027648; BAB55264.1; ..			
SO	SEQUENCE 804 AA: 90554 MW; 23C989DE603D208CD CRC64;			

Query Match	26.98;	Score 572.5;	DB 4;	Length 804;
Best Local Similarly	35.38;	Pred. No. 1.3e40;		
Matches 145;	Conservative 65;	Mismatches 158;	Indels 43;	Gaps 12

[illegible][illegible]

RESULT	3	
096SL5		
ID	096SL5	PRELIMINARY; PR1: 736 AA.
AC	096SL5;	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	CDNA FLJ147474 FIS, CLONE NT2RP4000051, WEAKLY SIMILAR TO	
DE	SYNAPTONEMAL COMPLEX PROTEIN SC65.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
OX	NCBI_Taxid=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	Nishikawa T., Nagai K., Sugeno S., Aotsuka S., Yoshikawa Y.,	
RA	Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu	
RA	Nakamura Y., Nagahari K., Masuno Y., Sasaki N.,	
RT	"NEDO human cDNA sequencing project."	
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.	
SR	EMBL; AK027680; BAB55291.1; ..	
SO	SEQUENCE 736 AA: 83411 MW; EA18193938FA685E CRC64;	

Query Match	26.88;	Score 571;	DB 4;	length 736;
Best Local Similarity	34.48;	Pred. No. 1.5e+40;		
Matches 145; Conservative	68;	Mismatches 168;	Indels 40;	Gaps 11;

```

OY      6 RGAHALALLLCVACALIRAGROQYERYSTRSPRDBELMPLESATYHNLADKISGEHMAESXG 65
Db      4 RALKLLTTLAVVAA--ASQAEVESEAGWGWTPDLL----FAEGTAAALARGDWPGVVL 56
OY      66 YLEISLRHLRLRDSSEAFCHNCSSA-----APQEPAGIASYPELRJFGGLLR 115
Db      57 SMERALSRBALRMLRLRCRTOCCADRPWELDRPMWSSPMAOSGAALRLRSPFGGLLR 116
OY      116 AHCLKRCQGLRPAFRQOSPREVLADQREBYKFLDPAFYAKNNLPKALIAAHTLKLH 175
Db      117 AACLRRC-LGPPAHSL--SEMELEBRKSPYUVOVAFAFKINKLEKAAVAAHTFVGN 173
OY      176 PDDEMKKNNMAYVUSLPGAEVY-IKDLETSSYSLFIRAVANGEMWRTSIDMETALP 234
Db      174 PEHHEMOQNLDTYOTMSGVKADKRDLETOHMEFRLGRLYSSEDRPOLAVHLEALQ 233
OY      235 DFFKAHYECIACGSGREIKDFK-----DFYLSIADHYVEVECKIQCSENL-TPYIGG 287
Db      234 EYFAVAYECCRALCGPYDYDDYNLLEYNADLFQALITDHYIOLVINCQNCVTETASHPSRE 293
OY      288 YPVEKFAVTWYHNYLOFAYUKLNDLKNAPACVSYLLFDONDKWMOQLVYUOYHNDRWGL 347
Db      294 KPFEDELFSHNYLOFAIYINIGNTQAVCEACKTYLLEFPNDEVANNOLAYTAAH-----L 348
OY      348 SDEHFO--PPREAVOFENNTYLOKELYDEFAKE---NIMDDGEVEVEYVDDLLELET 400
Db      349 GEEHTRSIGPRESAKEYORSELEKELLFPAYDVGFIPFDVPSMTPEEYIPKRLQEKOR 408
OY      401 S 401
Db      409 S 409

```


RESULT 4

O96SK8

ID O96SK8 PRELIMINARY; PRT: 736 AA.

AC O96SK8; 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DE CDNA FLJ14791 FIS. CLONE NT2RP4001064, WEAKLY SIMILAR TO

SYNATONEMAL COMPLEX PROTEIN SC65.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,

RA Nakamura Y., Nagahari K., Masubo Y., Sasaki N.,

RT "NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AK02697; BAB5305.1; -

SEQUENCE 736 AA; 83393 MW; EA1909828FAE685E CRC64;

Query Match 26.8%; Score 571; DB 4; Length 736;

Best Local Similarity 34.4%; Pred. No. 1.5e-40;

Matches 145; Conservative 68; Mismatches 168; Indels 40; Gaps 11;

QY 6 RGAALLLALCVACALRAGROYERSFRSPDELMPLSAVRHMLDKYSGEHWMSXG 65

DB 4 RALKLTLLAVVAA--ASQAEVSEKGMGVTPTDLL----FAEGTAAYARGDMPGVVL 56

QY 66 YLEISLRHLRLDSEAFCHRNCSA-----APQEPAGLASYPELRFGGLLR 115

DB 57 SMERALRSRALRALRLRCTQCAADFPWELDPDWSPPAASGAGALRDLSPFGGLLR 116

QY 116 AHLKRCOGSLPAPROSOPREVLADFORREPKFLQFAFKANNLPKATAAHTFLKH 175

DB 117 AALRLRC-LGPPAHSI--SEEMLEFRKSPYNYLQVAYFKINKLEKAAVAHAHTFVGN 173

QY 176 PDDEMKRNAAVYKSLPGAEDY- IKDLETSYSLSFLRAVRAYNGENWRTSITMETALP 234

DB 174 PEMEMOQNDLYQTMGSGVEADPKDLETPHMOEFLGRLYSEDPQCAVPHLEALQ 233

QY 235 DEFKATYECLACEGSREIKDFK-----DFYLSADHYVEVECKIQCEENL-TPYIGG 287

DB 234 EYFVAEEECALCEGPDYDGYNYLEYNADLFQAITDHYIQVINCCKONCVTELASHPSRE 293

QY 288 YPEKFPVATMYHILQFAVYKLNLDKNAAPCAVSYLLFDONDKYMOQLVYYOYHRDTWGL 347

DB 294 KPFEDELPSHYNYLQFAVYNYIGNTQAVCAKTYLLFFPNDEVNOMLAYAAM-----L 348

QY 348 SDEHFO--PRPVAQFENVNTLQKELYDFAKE---NIMDDGEVEVEVDLLELEET 400

DB 349 GEETHRSIGRESAKERYQSLLEKELLFAYDVFGIPYDPDSWTPPEVYIPRLQKOK 408

QY 401 S 401

DB 409 S 409

RESULT 5

O9HC86

ID O9HC86 PRELIMINARY; PRT: 736 AA.

AC O9HC86; 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE GROS1-L PROTEIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,

RA Nakamura Y., Nagahari K., Masubo Y., Sasaki N.,

RT "NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AK02697; BAB5305.1; -

SEQUENCE 736 AA; 83413 MW; 59D1205DB76ADB50 CRC64;

Query Match 26.6%; Score 567; DB 4; Length 736;

Best Local Similarity 34.2%; Pred. No. 3.4e-40;

Matches 144; Conservative 68; Mismatches 169; Indels 40; Gaps 11;

QY 6 RGAALLLALCVACALRAGROYERSFRSPDELMPLSAVRHMLDKYSGEHWMSXG 65

DB 4 RALKLTLLAVVAA--ASQAEVSEKGMGVTPTDLL----FAEGTAAYARGDMPGVVL 56

QY 66 YLEISLRHLRLDSEAFCHRNCSA-----APQEPAGLASYPELRFGGLLR 115

DB 57 SMERALRSRALRALRLRCTQCAADFPWELDPDWSPPAASGAGALRDLSPFGGLLR 116

QY 116 AHLKRCOGSLPAPROSOPREVLADFORREPKFLQFAFKANNLPKATAAHTFLKH 175

DB 117 AALRLRC-LGPPAHSI--SEEMLEFRKSPYNYLQVAYFKINKLEKAAVAHAHTFVGN 173

QY 176 PDDEMKRNAAVYKSLPGAEDY- IKDLETSYSLSFLRAVRAYNGENWRTSITMETALP 234

DB 174 PEMEMOQNDLYQTMGSGVEADPKDLETPHMOEFLGRLYSEDPQCAVPHLEALQ 233

QY 235 DEFKATYECLACEGSREIKDFK-----DFYLSADHYVEVECKIQCEENL-TPYIGG 287

DB 234 EYFVAEEECALCEGPDYDGYNYLEYNADLFQAITDHYIQVINCCKONCVTELASHPSRE 293

QY 288 YPEKFPVATMYHILQFAVYKLNLDKNAAPCAVSYLLFDONDKYMOQLVYYOYHRDTWGL 347

DB 294 KPFEDELPSHYNYLQFAVYNYIGNTQAVCAKTYLLFFPNDEVNOMLAYAAM-----L 348

QY 348 SDEHFO--PRPVAQFENVNTLQKELYDFAKE---NIMDDGEVEVEVDLLELEET 400

DB 349 GEETHRSIGRESAKERYQSLLEKELLFAYDVFGIPYDPDSWTPPEVYIPRLQKOK 408

QY 401 S 401

DB 409 S 409

RESULT 6

O9CWM5

ID O9CWM5 PRELIMINARY; PRT: 742 AA.

AC O9CWM5; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DE 2410024C15RIK PROTEIN.

GN GROS1 OR 2410024C15RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,

RA Nakamura Y., Nagahari K., Masubo Y., Sasaki N.,

RT "NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF097432; AAG31019.1; -

InterPro: IPR000886; ER_TARGET.

PROSITE: PS00014; ER_TARGET; UNKNOWN.1.

SEQUENCE 736 AA; 83413 MW; 59D1205DB76ADB50 CRC64;

Query Match 26.6%; Score 567; DB 4; Length 736;

Best Local Similarity 34.2%; Pred. No. 3.4e-40;

Matches 144; Conservative 68; Mismatches 169; Indels 40; Gaps 11;

QY 6 RGAALLLALCVACALRAGROYERSFRSPDELMPLSAVRHMLDKYSGEHWMSXG 65

DB 4 RALKLTLLAVVAA--ASQAEVSEKGMGVTPTDLL----FAEGTAAYARGDMPGVVL 56

QY 66 YLEISLRHLRLDSEAFCHRNCSA-----APQEPAGLASYPELRFGGLLR 115

DB 57 SMERALRSRALRALRLRCTQCAADFPWELDPDWSPPAASGAGALRDLSPFGGLLR 116

QY 116 AHLKRCOGSLPAPROSOPREVLADFORREPKFLQFAFKANNLPKATAAHTFLKH 175

DB 117 AALRLRC-LGPPAHSI--SEEMLEFRKSPYNYLQVAYFKINKLEKAAVAHAHTFVGN 173

QY 176 PDDEMKRNAAVYKSLPGAEDY- IKDLETSYSLSFLRAVRAYNGENWRTSITMETALP 234

DB 174 PEMEMOQNDLYQTMGSGVEADPKDLETPHMOEFLGRLYSEDPQCAVPHLEALQ 233

QY 235 DEFKATYECLACEGSREIKDFK-----DFYLSADHYVEVECKIQCEENL-TPYIGG 287

DB 234 EYFVAEEECALCEGPDYDGYNYLEYNADLFQAITDHYIQVINCCKONCVTELASHPSRE 293

QY 288 YPEKFPVATMYHILQFAVYKLNLDKNAAPCAVSYLLFDONDKYMOQLVYYOYHRDTWGL 347

DB 294 KPFEDELPSHYNYLQFAVYNYIGNTQAVCAKTYLLFFPNDEVNOMLAYAAM-----L 348

QY 348 SDEHFO--PRPVAQFENVNTLQKELYDFAKE---NIMDDGEVEVEVDLLELEET 400

DB 349 GEETHRSIGRESAKERYQSLLEKELLFAYDVFGIPYDPDSWTPPEVYIPRLQKOK 408

QY 401 S 401

DB 409 S 409

RESULT 6

O9CWM5

ID O9CWM5 PRELIMINARY; PRT: 742 AA.

AC O9CWM5; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DE 2410024C15RIK PROTEIN.

GN GROS1 OR 2410024C15RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,

RA Nakamura Y., Nagahari K., Masubo Y., Sasaki N.,

RT "NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF097432; AAG31019.1; -

InterPro: IPR000886; ER_TARGET.

PROSITE: PS00014; ER_TARGET; UNKNOWN.1.

SEQUENCE 736 AA; 83413 MW; 59D1205DB76ADB50 CRC64;

Query Match 26.6%; Score 567; DB 4; Length 736;

Best Local Similarity 34.2%; Pred. No. 3.4e-40;

Matches 144; Conservative 68; Mismatches 169; Indels 40; Gaps 11;

QY 6 RGAALLLALCVACALRAGROYERSFRSPDELMPLSAVRHMLDKYSGEHWMSXG 65

DB 4 RALKLTLLAVVAA--ASQAEVSEKGMGVTPTDLL----FAEGTAAYARGDMPGVVL 56

QY 66 YLEISLRHLRLDSEAFCHRNCSA-----APQEPAGLASYPELRFGGLLR 115

DB 57 SMERALRSRALRALRLRCTQCAADFPWELDPDWSPPAASGAGALRDLSPFGGLLR 116

QY 116 AHLKRCOGSLPAPROSOPREVLADFORREPKFLQFAFKANNLPKATAAHTFLKH 175

DB 117 AALRLRC-LGPPAHSI--SEEMLEFRKSPYNYLQVAYFKINKLEKAAVAHAHTFVGN 173

QY 176 PDDEMKRNAAVYKSLPGAEDY- IKDLETSYSLSFLRAVRAYNGENWRTSITMETALP 234

DB 174 PEMEMOQNDLYQTMGSGVEADPKDLETPHMOEFLGRLYSEDPQCAVPHLEALQ 233

QY 235 DEFKATYECLACEGSREIKDFK-----DFYLSADHYVEVECKIQCEENL-TPYIGG 287

DB 234 EYFVAEEECALCEGPDYDGYNYLEYNADLFQAITDHYIQVINCCKONCVTELASHPSRE 293

QY 288 YPEKFPVATMYHILQFAVYKLNLDKNAAPCAVSYLLFDONDKYMOQLVYYOYHRDTWGL 347

DB 294 KPFEDELPSHYNYLQFAVYNYIGNTQAVCAKTYLLFFPNDEVNOMLAYAAM-----L 348

QY 348 SDEHFO--PRPVAQFENVNTLQKELYDFAKE---NIMDDGEVEVEVDLLELEET 400

DB 349 GEETHRSIGRESAKERYQSLLEKELLFAYDVFGIPYDPDSWTPPEVYIPRLQKOK 408

QY 401 S 401

DB 409 S 409

RESULT 6

O9CWM5

ID O9CWM5 PRELIMINARY; PRT: 742 AA.

AC O9CWM5; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DE 2410024C15RIK PROTEIN.

GN GROS1 OR 2410024C15RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,

RA Nakamura Y., Nagahari K., Masubo Y., Sasaki N.,

RT "NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF097432; AAG31019.1; -

InterPro: IPR000886; ER_TARGET.

PROSITE: PS00014; ER_TARGET; UNKNOWN.1.

SEQUENCE 736 AA; 83413 MW; 59D1205DB76ADB50 CRC64;

Query Match 26.6%; Score 567; DB 4; Length 736;

Best Local Similarity 34.2%; Pred. No. 3.4e-40;

Matches 144; Conservative 68; Mismatches 169; Indels 40; Gaps 11;

QY 6 RGAALLLALCVACALRAGROYERSFRSPDELMPLSAVRHMLDKYSGEHWMSXG 65

DB 4 RALKLTLLAVVAA--ASQAEVSEKGMGVTPTDLL----FAEGTAAYARGDMPGVVL 56

QY 66 YLEISLRHLRLDSEAFCHRNCSA-----APQEPAGLASYPELRFGGLLR 115

DB 57 SMERALRSRALRALRLRCTQCAADFPWELDPDWSPPAASGAGALRDLSPFGGLLR 116

QY 116 AHLKRCOGSLPAPROSOPREVLADFORREPKFLQFAFKANNLPKATAAHTFLKH 175

DB 117 AALRLRC-LGPPAHSI--SEEMLEFRKSPYNYLQVAYFKINKLEKAAVAHAHTFVGN 173

QY 176 PDDEMKRNAAVYKSLPGAEDY- IKDLETSYSLSFLRAVRAYNGENWRTSITMETALP 234

DB 174 PEMEMOQNDLYQTMGSGVEADPKDLETPHMOEFLGRLYSEDPQCAVPHLEALQ 233

QY 235 DEFKATYECLACEGSREIKDFK-----DFYLSADHYVEVECKIQCEENL-TPYIGG 287

DB 234 EYFVAEEECALCEGPDYDGYNYLEYNADLFQAITDHYIQVINCCKONCVTELASHPSRE 293

QY 288 YPEKFPVATMYHILQFAVYKLNLDKNAAPCAVSYLLFDONDKYMOQLVYYOYHRDTWGL 347

DB 294 KPFEDELPSHYNYLQFAVYNYIGNTQAVCAKTYLLFFPNDEVNOMLAYAAM-----L 348

QY 348 SDEHFO--PRPVAQFENVNTLQKELYDFAKE---NIMDDGEVEVEVDLLELEET 400

DB 349 GEETHRSIGRESAKERYQSLLEKELLFAYDVFGIPYDPDSWTPPEVYIPRLQKOK 408

QY 401 S 401

DB 409 S 409


```

0Y 14 LILVACALRAGAOYERXSFSFPEDELMPLESAVRAHLADKYSGHMAESXGYLEISRL 73
Db 1 MVAVAAA--AASPAVAE-----SEPENVAAADLLIYAEGSTAAYAGDWMGVVLLNEMRALRS 54
0Y 74 HRLLDSEAFCHRNCSA----APOPE--PAA-----GLASYPERLRLGGLLRACHLCRLR 123
Db 55 RAALRALALRLCRTRKATELPPAHPDLDLGPASSLNHDPAAALHLDRIFGALLRRRAACLRLR 111
0Y 122 CKOGLPAFRQSOPSREVLADFORREPYKFLQFAYFAKANLLEKALIAAATFLPKHPDDEM 18
Db 115 C-LGPPSAHLL--SELDLEFNKRSPNYLQVAYFKINKLEKAVAAAHTEFVGNDPEHMEM 177
0Y 182 KRMAAYYSLSLGA--EDYIKOLETSYSLESLFRAVRAIVGEMWRKRSITDMELADEDFKRAF 244
Db 172 RONLDYITGMSGVKEEDDELEAKRPHMEFPLGVRLYSEEPRLAAYPHLEALAOEYFVAD 233
0Y 241 YECLACGSGREIKDFK-----PEYLSIAHYAVEVELECKTLOCEBNT--TPVITGGYPERKE 293
Db 232 BECRALCEGPPDYDGYANLIDYSAIDLFOAITHYIVQVYLSCKNOCYTELASHSREKPEDF 291
0Y 294 VATMYHYLOFAYYKILNDLKNAPCAVSAVSLIEDDNDKVMQOQLVYYOYHRDITWGLSDEH-- 351
Db 292 LPSHNYLQFAYYUNIGNYTQALIECAKTYLLLFEPNDEVMSQMLAYY-----TAVLGESEAS 344
0Y 352 --FORPEAVQFPNWTLOKEL-----YDPAKSINIMDDGEVGEVYVDLLLELEETS 401
Db 347 SISPEANQOEVRHRSLEKELFFAYVLDITFGIPLVDPOSWTPEEYVYPKLCKOKS 401

```

[illegible]

Db	216	RLISEKRGQEAHPHLEALQCYFYFADEBCRALCEGPRDYGYNYLNDYSALDPAITDHY	275
Qy	269	EVLBECKIOCEENTLTPYIGGYPVEKFWATMYHYLOFAVYKXKLNDKNAPCAVSYLLEFDON	327
Db	276	QVLMCKQVCVTELASHPRSREKPFEDFLPSHNYNLQFAVYNYNGNTQALIBECARITYLLEFPN	335
Qy	328	DKVMQONLVYQYIRDTWGLSDEN---EQPPEAVQFNVNTTLOKEL---YPAKEMIM	380
Db	336	DEVMQONLAVY-----TAMLGEESASSISPENAEKERRRNLLLEKELFFAYVDIFGIPAV	390
Qy	381	DDDEGEVVEYVDLLELEEMTS	401
Db	391	DPDSWTPEEVIPKRLQEQKS	411

RESULT	10			
Q9QZT7				
ID	Q9QZT7	PRELIMINARY;	PRT;	747 AA.
AC	Q9QZT7;			
DT	01-MAY-2000 (Tremblrel, 13, Created)			
DT	01-MAY-2000 (Tremblrel, 13, Last sequence update)			
DT	01-DEC-2001 (Tremblrel, 19, Last annotation update)			
DE	GROWTH SUPPRESSOR IL.			
CN	GROSL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CD1 ICR; TISSUE=FTBROBLAST;			
RX	MEDLINE=20406537; PubMed=10951563;			
RA	Kaul S.C., Sugihara T., Yoshida A., Nomura H., Wadhwa R.;			
RT	"Grosi, a potential growth suppressor on chromosome 1: its identity to			
RT	basement membrane-associated proteoglycan, leprecan.";			
RL	OncoGene 19:3576-3583(2000).			
DR	EMBL; AF165163; AAF04806.1; --			
DR	MED; MGI:1888921; Grosi.			
SO	SEQUENCE	747 AA;	84811 MW;	EBEAB2A62E824123 CRC64;

	Query Match	25.0%; Score 532.5; DB 11; Length 747;
	Best Local Similarity	35.4%; Pred. No. 3.2e-37;
	Matches 135; Conservative	57; Mismatches 134; Indels 35; Gaps 11;
OY	48 YRNAADKXSGEHWMSXGYLEISLRLHRLDRSEAFCHRNCSA-----APQ----	PEPA-- 97
Dd	39 YAEGTAAASRRDMDPEVVLNLMERALRSRAALRALRLRCRTTCACTELPMAPDDIDGDPPLS	98
OY	98 --AGIASPELRIRLGGLRLRAHCLCKRCKOGILPAFQOSOPSREVLADFORREPYKTLOP	155
Dd	99 QDPGAALAHDLRFEGAVLRRACCLRRC-LGPSSALL--SEELDIEFNKRSPNYNQVAY	155
OY	156 FKANLLPKAIATAAHFFLLKHPPDDEMKNNMYKSLPCAEIV-IKDLETKSYESIFRAY	214
Dd	156 FKINKLEAAVAARATFEFGNGNPEDMHMRONLDIYQMSGVKENDFDLAKPRMHEFRIGV	215
OY	215 RAYNGENNRITSITDMETALPDFKFAVECLAACEGREIKDF-----DFYLSTADHYV	268
Dd	216 RLYSSEKPOEAEPHLDEALOEYFVADCEBRALCESPDYDYGGYNYLDYSADLLEQAATTDDYV	275
OY	269 EYLEEKIQCEENL-TPVIGGVPEVEFVATMTHTLYOFATYYKLNDLKNAAPCAVSYLLFPQN	327
Dd	276 QVLNCKQCVCVELASHPSREKPFEEFLPSHSNLYLOFAFYNTGNATQALECAKYTYLLFPNP	335
OY	328 DKVMQONLVVVYQYHHDWTMGSLSDER---PQPPEAQVFNVNTTLQEL-----YDAFKERNI	380
Dd	336 DEVMHQNLAYI-----TMLGEESKASSSPKNEMNEYRRPNLKEKELLFPAYDIGFIIFY	390
OY	381 DDDEGEVEVEVDLLILEETS	401
Dd	391 DPDSVTPEEVIPKRIQEOKS	411

Query Match 7.5%; Score 159.5; DB 4; Length 551;
 Best Local Similarity 24.8%; Pred. No. 2.4e-05;
 Matches 54; Conservative 43; Mismatches 92; Indels 29; Gaps 7;

```

QY 181 MKRNAYKSLPGAE-DYIKDLETKSYESLFIKAVRAVNGENWRTSTDMELALPDFKA 239
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
Db 5 MREDMAKXRRMSGVRPOSFRDLTPPHWAAYDTGLGROEAGLALPRLEALQGSIAQ 64
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
QY 240 FYECLACEGSRREIKDFD-----FYLSIADHYVEVLECKIOC-EENLTPVIGY 288
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
Db 65 MESCRADCEGPEEQGAEFEEDGAASQGLYEALAGHWIQVLCRQRCVGEAATPGRSF 124
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
QY 289 PVEKFAATMYHYLOFAYYKLNLDKNAPCAVSYLLFDONDKVMQONLYYQYHRDTWGLS 348
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
Db 125 FVPDFLPNQLRLRLHEAHQVGNLSQAIENVLVSLLFYPEDEAKRALNQQAQLG----- 179
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
QY 349 DEHQPRP-----EAVQFFNVTTL--QKELYDFAKENI 379
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
Db 180 ---EPRGLGPREDIQRFILRSLSGKROLY-YAMEHL 212
      ||| | : : : ||| | : : : : : : : : : : : : : : : : : :
  
```

Search completed: June 8, 2002, 02:16:59
 Job time: 484 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2002, 00:10:15 ; Search time 32.49 Seconds

(Without alignments)
301,467 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130
Sequence: 1 MEPRGGAALLALCVACA.....DDEGEVVEYDLELEFETS 401

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.5	4.4	816	2	US-08-785-310A-8
2	94.5	4.4	816	2	US-08-816-693A-53
3	94.5	4.4	816	3	US-08-885-291-53
4	94.5	4.4	816	4	US-09-496-672-53
5	92.5	4.3	806	1	US-08-451-715A-6
6	91.5	4.3	486	1	US-08-328-256-12
7	90	4.2	531	2	US-08-923-536A-12
8	89.5	4.2	855	2	US-08-816-693A-2
9	89.5	4.2	855	3	US-08-885-291-2
10	89.5	4.2	855	4	US-09-496-672-2
11	87	4.1	548	1	US-08-333-358-2
12	87	4.1	548	1	US-08-463-694-2
13	87	4.1	548	1	US-08-694-501-2
14	87	4.1	556	4	US-09-106-194-3
15	87	4.1	1708	1	US-08-493-092-2
16	87	4.1	1708	1	US-08-508-836A-2
17	87	4.1	3056	1	US-08-508-836A-8
18	87	4.1	3056	2	US-08-629-001A-3
19	87	4.1	3056	2	US-08-874-266-2
20	87	4.1	3056	4	US-08-642-274D-3
21	87	4.1	3056	4	US-08-952-127-3
22	87	4.1	3056	4	US-08-952-014C-3
23	86	4.0	590	2	US-08-980-060-2
24	86	4.0	590	2	US-09-307-185-2
25	85.5	4.0	846	3	US-08-885-291-55
26	85.5	4.0	846	3	US-09-107-847-2
27	85.5	4.0	846	4	US-09-496-672-55

28	84.5	4.0	1513	5	PCT-US93-03076-2	Sequence 2, Appl
29	84	3.9	557	1	US-08-328-256-10	Sequence 10, Appl
30	84	3.9	747	2	US-08-816-693A-51	Sequence 51, Appl
31	84	3.9	747	3	US-08-885-291-51	Sequence 51, Appl
32	84	3.9	747	4	US-09-496-672-51	Sequence 14, Appl
33	83.5	3.9	399	4	US-09-347-801-14	Sequence 20, Appl
34	83.5	3.9	405	4	US-09-347-801-20	Sequence 2, Appl
35	83.5	3.9	688	2	US-09-016-000-2	Sequence 11, Appl
36	83	3.9	434	1	US-08-328-256-11	Sequence 2, Appl
37	83	3.9	456	2	US-08-307-588-2	Sequence 2, Appl
38	83	3.9	557	1	US-08-471-454-2	Sequence 2, Appl
39	83	3.9	557	2	US-08-466-974-2	Sequence 2, Appl
40	83	3.9	557	2	US-08-471-453-2	Sequence 4, Appl
41	83	3.9	557	2	US-08-307-588-4	Sequence 4, Appl
42	83	3.9	625	2	US-08-980-060-4	Sequence 4, Appl
43	83	3.9	625	4	US-09-307-185-4	Sequence 4, Appl
44	83	3.9	824	2	US-08-785-310A-7	Sequence 7, Appl
45	82.5	3.9	608	4	US-09-315-793-22	Sequence 22, Appl

ALIGNMENTS

RESULT

1
US-08-785-310A-8
; Sequence 8, Application US/08785310A
; Patent No. 5840532
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; TITLE OF INVENTION: Neuronal PAS Domain Protein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,310A
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-785-310A-8

Query Match

Best local Similarity 4.4%; Score 94.5; DB 2; Length 816;
Matches 62; Conservative 41; Mismatches 123; Indels 87; Gaps 14;

QY 126 LPPEFOSQPREVLADQREPKFLQFAVFKANNLPKALTAHAHTLTKPPDDMK--R 183
DB 120 LPADVMDNLNLPPEQSEVYKIL-----SSHMLVTQSPSPPEFLKSDN 164

[illegible]

```

RESULT 3
US-08-885-291-53
: Sequence 53, Application US/08885291A
: Patent No. 6057125
: GENERAL INFORMATION:
: APPLICANT: Takahashi, Joseph S.
: APPLICANT: Turek, Fred W.
: APPLICANT: Pinto, Lawrence H.
: TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
: FILE REFERENCE: 0280-5
: CURRENT APPLICATION NUMBER: US/08/885, 291A
: CURRENT FILING DATE: 1997-06-30
: EARLIER APPLICATION NUMBER: 08/816, 693
: EARLIER FILING DATE: 1997-03-13
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 53
: LENGTH: 816
: TYPE: prt
: ORGANISM: Mus musculus
US-08-885-291-53

```

Query Match	4.4%	Score 94.5	DB 3	Length 816
Best Local Similarity	19.8%	Pred. No. 0.46		
Matches	62	Conservative	41	Mismatches 123; Indels 87; Gaps 14

QY	126	LPAFROSQSPREVIADFQRRPEPKFLQRAYFANNLPKAIAAHFFLLKHPDDEMK--R	183
Db	120	LPADWMDONLFLFPEOSEVYKTL-----SSHMLVTDSPSPFLKSDN	164
QY	184	NMAVYKSLPGAEDYIKDLETKSYEL-FLRAYANG-----ENMRTS-----	225
Db	165	DLEFCHL-LRGSINPKPEPFYEYIKVGNKRSTANNPSPSCNGFDNLSKPCHPCLK	222
QY	226	---ITDMELAPDFKFAFYECIAACEGSRREIKDKDFYLSTADHYVEYLECK-IOCEEN	280
Db	223	DVCFATATVRLATPQLKEH-----CVADPELPEEFTSRH-----SLEWKFLDHR	267
QY	281	LTPYVGIGFVPEKFAVTMTHYLQFAYYKINDLKNAPCAVSYLLFPQNDK-----	329
Db	268	APPIIGYLPFEVLGTSGN-----YHNIIDDELLARCHOHMDQFKSGSCCYRFLTKGOQ	322

```

-      : | ||| - |   : | :||
Db    323 WIMIQTHYYITHQ--WNSKPEFVCTHSVSADYVERRQ-----ELATEDPPTTEAM 374
Oy    387 --VEEYVDLLE     396
       ! : | || 
Db    375 HPSAVKERDSSLE 387

```

RESULT ^A

US-09-496-672-53


```
; Sequence 53, Application US/09496672
; Patent No. 6291429
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/09/496,672
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 08/885,291
; PRIOR FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 08/816,693
; PRIOR FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 53
; LENGTH: 816
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-496-672-53
```

```
Query Match 4.4%; Score 94.5; DB 4; Length 816;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 62; Conservative 41; Mismatches 123; Indels 87; Gaps 14;
```

```
QY 126 LPAFROSQPSREVLADFORREPKFLQFAFKANNLPKAIAMHTLPLKIPDDMMK--R 183
Db 120 LPAVDMDONLNLPEQSHSEYKIL-----SSHMLVTSPSEFLKSDN 164
QY 184 NMAVYKSLPGEADYIKDLFKSYESL-FIRAVRAYNG-----ENWRTS----- 225
Db 165 DLEFYCHL-LRSLNPKPEPTYEIKFVGNFNSYNNVPSPSCNGPDNTLSRCHVPLGK 222
QY 226 ----TDMELALDPFKAFYECLAGEGREIKDFYLSIADHYVEVLECK-IOCEEN 280
Db 223 DVCFIATVRLATPOFLKEM-----CVADRLPEFTSRH-----SLEWKFLEFDR 267
QY 281 LTVYIGYIPREKVATMYHNLQAFAYKLNLDKNAAPCAVSYLLFDONDK----- 329
Db 268 APPIGLPEVLGTSGYN-----YYHIDLELLARCHOHMOFGKSCCYRFLTKGQO 322
QY 330 -WQOUNLVYYQYHRDTGSLDEHQRPRPAVQFFNVTTLOKELYDFAKENIMDDGE-- 386
Db 323 WIMLOHHYITTHQ--WNSRPEFTVCTHSVYSTADYRVERKQ-----ELALDDPTLEAM 374
QY 387 ---VEYVDLLE 396
Db 375 HPSAVKEKDSLE 387
```

```
RESULT 5
US-08-451-715A-6
; Sequence 6, Application US/08451715A
; Patent No. 5801013
; GENERAL INFORMATION:
```

```
; APPLICANT: Tao, Jianshi
; APPLICANT: Qul, Yan
; APPLICANT: Houman, Fariba
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising Same
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,715A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOK, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI94-25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 806 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-451-715A-6
```

```
Query Match 4.3%; Score 92.5; DB 1; Length 806;
Best Local Similarity 18.6%; Pred. No. 0.73;
Matches 57; Conservative 41; Mismatches 95; Indels 113; Gaps 12;
```

```
QY 177 DDEMKRNM-AVYKSLPG-AEDYIKDLTKSYSLFIRAVRAYNGENWRTSTDM----- 230
Db 176 DTEVQKELQYVYKILITNYAEELIKDELTE-----NHMPQVLTQKNNI 221
QY 231 -LALDPFKAYECLANQESREIKDFD-----FYLSTADHY-----VEV 270
Db 222 GKSIQLOFGRIADDECLKACNGIOELEVTTRADTIYGVYIAIAPHLVEHAIRVSO 281
QY 271 LECKI-----QCEENLTPVIGYVPEKRV----- 294
Db 282 EDSKIITKILNTQORALEKKAFLGVYAIHPLTKOKIPVWVANFALNAYSGALMGVP 341
QY 295 -----ATMYHY-----LOPAYKKLNLDKNAAPCAVSYLLFDONDXVMOONLV 336
Db 342 ACDEPDEFANLYHIPKITYQSPONLPHKKEVLNKGEMS-----DLSSEVAREOIL 395
QY 337 YV-----QYHRDTGSLDEHQRPRPAVQFFNVTTL--OKELYDFAKENIMDD 383
Db 396 AYFEKENLGRVINYRLQDWGVSRORYWGAIPDHICNCGIVPETQLPVLDPEDIVDG 455
QY 384 EGEVVE 389
Db 456 EGNPLE 461
```

```
RESULT 6
US-08-328-256-12
```

```
; Sequence 12, Application US/08328256
; Patent No. 5643749
; GENERAL INFORMATION:
```

```
; APPLICANT: REVEL, Michel
; APPLICANT: ABRAMOVICH, Carolina
; APPLICANT: RATOVIJSKI, Edward
; TITLE OF INVENTION: SOLUBLE INTERFERON ALPHA-RECEPTOR, ITS
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,256
; FILING DATE: 24-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 107378
; FILING DATE: 24-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: REVEL-13
; REFERENCE/DOCKET NUMBER: 25,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 496 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-328-256-12

```

```

Query Match          4.3%; Score 91.5; DB 1; Length 496;
Best Local Similarity 20.3%; Pred. No. 0.44;
Matches 84; Conservative 62; Mismatches 157; Indels 111; Gaps 22;

```

```

OY 28 YERYSFSPDELMPLSAYRHALDK-----YSEHMAES 63
    ||| | | | | | | | | | | | | | | | | | | |
DB 115 YEVSFEPFRKAQIGPPE-VHLEAEADKAIYHISPGTKDSVMALDGLSFYSLIMKNS 173
OY 64 XGY---LEISLRHLR---LRDSEAFCHRNCSAAPQPEPAALASYPELRGLRRHC 118
    ||| | | | | | | | | | | | | | | | | | | |
DB 174 SGVEERINITSRHITYKLSPETTYCLK-----VKAALLTSMTK-----GYSPVHC 220
OY 119 LKRCQKQ-LPAFROSOPSREVLADFORREPKFLQFAFYKANNLPKAIANAATFELKHPD 177
    ||| | | | | | | | | | | | | | | | | | | |
DB 221 IKTVENELP-----PRENIEVSQONN--YVLKMDTYA-NMTFQOMLHAFKRRPG 271
OY 178 DEMAKRNMYKSLPGADYIKDLETYSLSL-----IRAVRAVNGENMTSTIDMELA 232
    ||| | | | | | | | | | | | | | | | | | | |
DB 272 NHLVK---WKQIPDCN-VKTCQVFPQWFOKGYLLRVOASDGN--TSFMSSEIK 323
OY 233 LPDFKAFYECLACEGSREIKDFKDFYLSADHYVEVLECKICEENLTPVIGSYPYEK 292
    ||| | | | | | | | | | | | | | | | | | | |
DB 324 FDTETQAF-LLPVENIRSLSDSFHYIGAPK-----QSGNTFYIDYPL-- 367
OY 293 FVATMYHYLQFAYYKLANDKNAAPCAVSYLEFDQNDKVMQNL---VYQYHRDWTGLS 348
    ||| | | | | | | | | | | | | | | | | | | |
DB 368 -----IYEIIYEWNTS--NAERKIIKKTDTVTNKLPLVYCVKAR----- 407
OY 349 DEHFQRPPEAVOFFNVTTLQKELYDPFAKENIMDDDEGEVEYVDLLEETS 401
    ||| | | | | | | | | | | | | | | | | | | |
DB 408 -AHTMDESDAVCEYFSEQPLKNLLSTSEQI---EKCFIENISTATVETN 457

```

```

RESULT 7
US-08-923-536A-12
; Sequence 12, Application US/08923536A
; Patent No. 5965426
; GENERAL INFORMATION:
; APPLICANT: SAKAI, Yasuyoshi
; APPLICANT: KATO, No. 5965426uo
; APPLICANT: SHIBANO, Yuji
; TITLE OF INVENTION: PROTEIN DISULFIDE ISOMERASE GENE DERIVED
; TITLE OF INVENTION: FROM STRAIN OF METHYLOTROPIC YEAST
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIAS, L.L.P.

```

```

; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,536A
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-234287
; FILING DATE: 04-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-317
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-923-536A-12

```

```

Query Match          4.2%; Score 90; DB 2; Length 531;
Best Local Similarity 19.3%; Pred. No. 0.7;
Matches 79; Conservative 67; Mismatches 183; Indels 80; Gaps 16;

```

```

OY 9 AALLALLCVACALRAGR-----AQERYSFSPDELMPLSAYRHALDKYSGEHW 60
    ||| | | | | | | | | | | | | | | | | | | |
DB 10 ATILACTLVVADGGAIASPSAVVKTADSFESFKENPLVY-----AEFF 57
OY 61 AESGYL-----EISLRLHRLRSEAFCHRNCSAAPQPEPAALASYPELRGLRR 115
    ||| | | | | | | | | | | | | | | | | | | |
DB 58 APWCCHCKRRLGPEFOVADKLVKNDIRLAQIDCTEEKDLCSSGICKPTLKVFRGYENE 117
OY 116 A--HCLRCQKQGLPAF--ROSOPSREVLADFORREPKFLQFAFYKANNLPKAIANAATF 171
    ||| | | | | | | | | | | | | | | | | | | |
DB 118 PSDVAGQRTSDSISLWVWKGSTPPVSIVDDLSQIED-----TIKSNPQVFIO----- 165
OY 172 LKHPDDEMAKRNMYKSLPGADYIKDLETYS--YESLFIKAVRAVNGENMTSTIDM 229
    ||| | | | | | | | | | | | | | | | | | | |
DB 172 LKHPDDEMAKRNMYKSLPGADYIKDLETYS--YESLFIKAVRAVNGENMTSTIDM 229
OY 166 VLPKSGSKSVENGSTFEIANGLRDNYSFISTSTFSSSKYIKIKSDTPSYLLFRPNE 225
    ||| | | | | | | | | | | | | | | | | | | |
DB 230 ELALPDEFKAFYECLACEGSREIKDFKDFYLSADHYVEVLECKICEENLTPVIGSYP 289
    ||| | | | | | | | | | | | | | | | | | | |
DB 226 ELSDASITK-----FDEIID-----THLIEFLANESK-----PLFEBMD 259
OY 290 VEKFAVTHYHYLQFAYYKLANDL--KNAAPCAVSYLEFDQNDKVMQNLVYQYHRDWTGL 347
    ||| | | | | | | | | | | | | | | | | | | |
DB 260 GSSFQSYWEMKLPVAYFYFNISEKDAVSDAISKLATTHRGKVFVGLDSKSY-----GL 314
OY 348 SDEHFQRPPEAVOFFNVTTLQKEL-YDPFAKENIMDDDEGEVEYVDL 395
    ||| | | | | | | | | | | | | | | | | | | |
DB 315 HAKNINMK-EEFPLFAIHDIALELKYSIDSKPLDNK--LIPKVEDEVFV 360

```

```

RESULT 8
US-08-816-693A-2
; Sequence 2, Application US/08816693A
; Patent No. 5874241
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S
; APPLICANT: Turek, Fred W

```

```

; APPLICANT: Pinto, Lawrence H
; TITLE OF INVENTION: Clock Gene and Gene Product
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rockey, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,693A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5874241thrup, Thomas E
; REGISTRATION NUMBER: 33,268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-816-693A-2

```

```

Query Match          4.2%; Score 89.5; DB 2; Length 855;
Best Local Similarity 20.3%; Pred. No. 1.7;
Matches 58; Conservative 36; Mismatches 111; Indels 81; Gaps 11;

QY 110 GGLIRRAHCLKCKOGIPAFROSOPSRREVLADEFQRRPYFLOFAFYKANL-PKAIAAA 166
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 129 GSIIYSESVTSLEHLPDSLDVDSINFPEGSHSEYKILSTHLESOSLPEYLSK 188
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 169 H-----TFLLKHPDDEMKRMAYKSLP-----GAEDYIKDLETKSYES- 208
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 189 NQLEFCCHMLRGITDPKEPSTYEYVRIGNFKSLTSYSTHNGFEETIQRTHRPSYEDR 248
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 209 -LPIRAVRAYNGENWRTSTDMELALDPFKAFYECLAGE--GSREIKDFKDYLSIAD 265
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 249 VCFVATVR-----LATPOFIKEMCTVEEPNEEFTSRHSLEWKFLFL--D 290
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 266 HYVEVLECKIQCEENLTPVIGYVEKEFVATMTHYLOFAYYKLDLKNAPCAVSYLLFD 325
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 291 H-----RAPPIIGYLPFEVLGTSYD-----YHVDLENLAKCHEHLMQYG 332
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 326 QNDK-----VMOQNLVYYOYHRDPTWGLSDEHFQRPRAV 359
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 333 GKSCSYRFLTKGQGWIMLOTHTYITTHQ-----WMSRPEFI 369
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 9
US-08-885-291-2
; Sequence 2, Application US/08885291A
; Patent No. 6057125
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/08/885,291A
; CURRENT FILING DATE: 1997-06-30
; EARLIER APPLICATION NUMBER: 08/816,693
; EARLIER FILING DATE: 1997-03-13

```

```

; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-885-291-2

```

```

Query Match          4.2%; Score 89.5; DB 3; Length 855;
Best Local Similarity 20.3%; Pred. No. 1.7;
Matches 58; Conservative 36; Mismatches 111; Indels 81; Gaps 11;

QY 110 GGLIRRAHCLKCKOGIPAFROSOPSRREVLADEFQRRPYFLOFAFYKANL-PKAIAAA 166
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 129 GSIIYSESVTSLEHLPDSLDVDSINFPEGSHSEYKILSTHLESOSLPEYLSK 188
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 169 H-----TFLLKHPDDEMKRMAYKSLP-----GAEDYIKDLETKSYES- 208
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 189 NQLEFCCHMLRGITDPKEPSTYEYVRIGNFKSLTSYSTHNGFEETIQRTHRPSYEDR 248
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 209 -LPIRAVRAYNGENWRTSTDMELALDPFKAFYECLAGE--GSREIKDFKDYLSIAD 265
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 249 VCFVATVR-----LATPOFIKEMCTVEEPNEEFTSRHSLEWKFLFL--D 290
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 266 HYVEVLECKIQCEENLTPVIGYVEKEFVATMTHYLOFAYYKLDLKNAPCAVSYLLFD 325
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 291 H-----RAPPIIGYLPFEVLGTSYD-----YHVDLENLAKCHEHLMQYG 332
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 326 QNDK-----VMOQNLVYYOYHRDPTWGLSDEHFQRPRAV 359
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 333 GKSCSYRFLTKGQGWIMLOTHTYITTHQ-----WMSRPEFI 369
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 10
US-09-496-672-2
; Sequence 2, Application US/09496672
; Patent No. 6291429
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/09/496,672
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 08/885,291
; PRIOR FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 08/816,693
; PRIOR FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-496-672-2

```

```

Query Match          4.2%; Score 89.5; DB 4; Length 855;
Best Local Similarity 20.3%; Pred. No. 1.7;
Matches 58; Conservative 36; Mismatches 111; Indels 81; Gaps 11;

QY 110 GGLIRRAHCLKCKOGIPAFROSOPSRREVLADEFQRRPYFLOFAFYKANL-PKAIAAA 166
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 129 GSIIYSESVTSLEHLPDSLDVDSINFPEGSHSEYKILSTHLESOSLPEYLSK 188
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 169 H-----TFLLKHPDDEMKRMAYKSLP-----GAEDYIKDLETKSYES- 208
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 189 NQLEFCCHMLRGITDPKEPSTYEYVRIGNFKSLTSYSTHNGFEETIQRTHRPSYEDR 248
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 209 -LPIRAVRAYNGENWRTSTDMELALDPFKAFYECLAGE--GSREIKDFKDYLSIAD 265
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Db 249 VCFATVR-----LATPOFIKMCIVEEPPNEEFSTRHSLEKFLFL---D 290
Qy 266 HVEVELECKIOCEENTLTVIGGYVEKEFVATWYHYLOFAYKLLDKNAAPCAVSLLFD 325
Db 291 H-----RAPPIGLPFEVIGTSGYD-----YIVVDLEMLAKHEHLMQYG 332
Qy 326 ONDK-----VMOQNLVYQYHBDTWGLSDEHPDPEAV 359
Db 333 GKSCYVRFELTKGOQWIMLQTHYITVHQ-----WNSRPEFI 369

RESULT 11
US-08-333-358-2
; Sequence 2, Application US/08333358
; Patent No. 5571696
; GENERAL INFORMATION:
; APPLICANT: EVANS Ph.D., RONALD M.
; APPLICANT: MANGELSDORF Ph.D., DAVID J.
; APPLICANT: ONG MS., ESTELITA S.
; APPLICANT: ORO Ph.D., ANTHONY E.
; APPLICANT: BORMEYER Ph.D., UWE K.
; APPLICANT: GIGIERE Ph.D., VINCENT NMN
; APPLICANT: YAO Mr., TSO-PANG NMN
; TITLE OF INVENTION: NOVEL RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: US
; ZIP: 90071-2921
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,358
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/761,068
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-333-358-2

Query Match 4.18; Score 87; DB 1; Length 548;
Best Local Similarity 20.28; Pred. No. 1.5;
Matches 93; Conservative 48; Mismatches 156; Indels 164; Gaps 21;

Qy 16 CVAC-----ALRAGR-----AQYERYSFRSFPD-----ELMPLISAYRNA--- 51
Db 158 CLAVGMSRDYKVFGRMSKKORDSLAEVQKHKRMQOQDHOQOPCEAPRPIPTVYISANG 217
Qy 52 -----LDKISGEHWAE-----SXGYLEISLRHLRLLDSEAFCHRNCSAAPQ-- 94
Db 218 LTELHDDLSNYIDGHTPEGSKADSAVSFYLDIQSPDQSGDINGI-----KPEPIC 270
Qy 95 --EPAAGIASY-----PELRLEGGILRAHCLKRC---KQGLAPFRQSQPS 135

Db 271 DYTASGFEPPYCSFTNGETSPTVSMAELEHLAQNISKSH-LETQYLRKEELQOITWQTF 329
Qy 126 REVLADEFORREPKYKLOFAVFAKNNLPKAIAAHFTLLKHPDDEMKRNMAVYKSLPG-- 193
Db 330 OEIENYONKOREVWMOJCAKIT-----EAIQYVEFAKRIDQFM 370
Qy 194 ---AEDYIKDLETKSYESLFIKRAVAYNGENWRSTITMELALPDFFKAFYECJLAACG- 249
Db 371 ELQNDQIYVLLKAGSLEVFIFMCRAPDSQ--NTVYFGKAYASPDVFKSL-----GCEDF 424
Qy 250 -----SREIKDFKDFYLSADH-----YVEVELECKIOCEENTLTVI 285
Db 425 ISFVEFEKSLCSMHLTEDELALFSAFVLMGADRSMLOEKVIEKLOKIO----- 475
Qy 286 GGYVEKEFVATWYHYLOFAYKLLDKNAAPCAVS---YLLFDQNDKWOONLVYQYHR 342
Db 476 -----LALQHVLO--KNHREDGILTKLICKYSTLALCGRHEKLMARKAIYPDIVR 525
Qy 343 DTWGLSDEHFOPRPEAVQFVNTYLLQKELYDFAKENIMDD 383
Db 526 -----LHFPP-----LYKELTSEFEPAQID 547

RESULT 12
US-08-463-694-2
; Sequence 2, Application US/08463694
; Patent No. 5696233
; GENERAL INFORMATION:
; APPLICANT: EVANS Ph.D., RONALD M.
; APPLICANT: MANGELSDORF Ph.D., DAVID J.
; APPLICANT: ONG MS., ESTELITA S.
; APPLICANT: ORO Ph.D., ANTHONY E.
; APPLICANT: BORMEYER Ph.D., UWE K.
; APPLICANT: GIGIERE Ph.D., VINCENT NMN
; APPLICANT: YAO Mr., TSO-PANG NMN
; TITLE OF INVENTION: NOVEL RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: US
; ZIP: 90071-2921
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,694
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/761,068
; FILING DATE: 17-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-694-2

```

NAME: Rafter Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/DOCKET NUMBER: P-1 8936
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-501-2

Query Match      4.1%; Score 87; DB 1; Length 548;
Best Local Similarity 20.2%; Pred. No. 1.5;
Matches 93; Conservative 48; Mismatches 156; Indels 164; Gaps 21;

QY    16 CVAC-----ALRAGR-----AQRYSFRSPRPD-----ELMPLESAYRNA--- 51
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    158 CLAVGMSNDAYKFGFMSKKORDSLYAEVQKHRRMQQQOQHDDHQOQPEAGPLPTFTYNISANG 217

QY    52 -----LDKSGEHMAE-----SKGYEISRHLRLHRLDSAFCHRCNSAAPDP-- 94
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    218 LTELHDDLSNYIDGHTHPGSKADSASVSSFYDIOPSPDOSGDINDGI-----KPEPIC 270

QY    95 --EPRAGLASY-----PRLRLPGGLLRRAHCLKRC--KOGLPAPRQSOPS 135
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    271 DYTASGFPPYCSTNGTSETSPVSMAELEHIAONISKSH-LETCOYLREELIOQTITWOTFL 329

QY    136 REVLADEFORREPYPKLOFAYPEKANNLPRKAIAAHFTLLKHBDDEMKRMMAVYKSLPG-- 193
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    330 QEEIENQNKRQREVWMOCAIKIT-----EAIVVEPAKRKDIDGM-- 370

QY    194 ---AEDYKDELTSYSSELFTRAVRAYGENWRKSITMELALPFFFAFYECIACBS- 249
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    371 ELCONDQIVLTKAGSLLEVFIIMCRAPFSQSN-NIVYFDGKYASPVPFKSL----GCEDF 424

QY    250 -----SREIKDFKDFYLTADH-----YVEVLCKICCEENLTPIV 285
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    425 ISVFPERGKSLCSMHLEDETALFSAFVLMASADRSMLOEKVKYTELOOKIO----- 475

QY    286 GGYPEVEKFVATMYHYLOFAYYKLNDLKNAAPCAVS---YLLFDQNDKWQOOLVYYQYHR 342
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    476 -----LALQHVLO-KNHRREDGITLKLICKYSTLRALCGRHTEKMLAFRAIPDLIVR 525

QY    343 DTWGLSDEHFQRPPEAVQFNVTTLQKLEYDPAKENIMDD 383
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    526 -----LHFP-----LYKELFTSEEPEAMQID 547

RESULT 14
US-09-106-194-3
Sequence 3, Application US/09106194
Patent No. 6262234
GENERAL INFORMATION:
APPLICANT: Holloway, James
APPLICANT: Jelinek, Laura
APPLICANT: Durnam, Diane
APPLICANT: Blumberg, Hal
TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR POLYPEPTIDE
TITLE OF INVENTION: ZPPAR4
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
```

```

1  COMPUTER: IBM Compatible
2  OPERATING SYSTEM: DOS
3  SOFTWARE: PASTSEQ for Windows Version 2.0
4  CURRENT APPLICATION DATA:
5  APPLICATION NUMBER: US/09/106.194
6  FILING DATE:
7  CLASSIFICATION:
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER:
10 FILING DATE:
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Leitch, Debra K
13 REGISTRATION NUMBER: 32,619
14 REFERENCE/DOCKET NUMBER: 96-11
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 206-442-6674
17 TELEFAX: 206-442-6678
18
19 TELEX:
20
21 INFORMATION FOR SEQ ID NO: 3:
22
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 556 amino acids
25 TYPE: amino acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28
29 MOLECULE TYPE: peptide
30
31 US-09-106-194-3

```

```

1  TITLE OF INVENTION:  Ataxia-Telangiectasia Gene
2  NUMBER OF SEQUENCES:  7
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  Reising, Ethlington, Barnard & Perry
5  STREET:  P.O. Box 4390
6  CITY:  Troy
7  STATE:  Michigan
8  COUNTRY:  US
9  ZIP:  48099
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE:  Floppy disk
12 OPERATING SYSTEM:  PC-DOS/MS-DOS
13 SOFTWARE:  PatentIn Release #1.0, Version #1.30
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER:  US/08/493,092
16 FILING DATE:
17 CLASSIFICATION:  435
18 ATTORNEY/AGENT INFORMATION:
19 NAME:  Kohn, Kenneth I.
20 REGISTRATION NUMBER:  30,955
21 REFERENCE/DOCKET NUMBER:  P-310 (TAY)
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE:  (810) 689-3500
24 TELEFAX:  (810) 689-4071
25 INFORMATION FOR SEQ ID NO:  2:
26 SEQUENCE CHARACTERISTICS:
27     LENGTH:  1708 amino acids
28     TYPE:  amino acid
29     STRANDEDNESS:  single
30     TOPOLOGY:  linear
31
32 US-08-493-092-2

```

Job time: 714 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2002, 00:14:20 ; Search time 49.19 Seconds

(without alignments)
783.326 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130

Sequence: 1 MEPCRGCAALLALICVACA.....DDEGEVEVEYVDLLEETS 401

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1019	47.8	431	2	A56822 synaptonemal compl
2	106	5.0	1345	2	A64430 DNA-directed RNA p
3	104	4.9	317	2	B95113 competence protein
4	104	4.9	317	2	A97982 competence protein
5	104	4.9	853	2	H70470 hypothetical prote
6	98.5	4.6	1575	2	G83905 conserved hypothet
7	98	4.6	376	2	T32009 hypothetical prote
8	98	4.6	484	2	E70031 hypothetical prote
9	97.5	4.6	910	2	T41331 proline dehydrogen
10	97.5	4.6	1162	2	C81297 proline dehydrogen
11	96	4.5	788	2	A71076 conserved hypothet
12	95.5	4.5	240	2	E90139 conserved hypothet
13	95	4.5	412	2	H65011 probable transamin
14	94.5	4.4	292	2	S64361 hypothetical prote
15	94	4.4	435	2	B81279 probable fibronect
16	94	4.4	922	2	S37871 hypothetical prote
17	93	4.4	422	2	AD2191 hypothetical prote
18	93	4.4	433	2	D91111 probable adherence
19	93	4.4	433	2	H85956 probable cytooxin
20	93	4.4	434	2	B97018 xre family DNA-bin
21	93	4.4	1101	2	AB2481 hypothetical prote
22	92.5	4.3	729	2	E81410 transcription regu
23	92	4.3	411	2	E83056 transaminase (EC 2
24	92	4.3	534	2	T41081 hypothetical prote
25	92	4.3	1132	2	T44001 major DNA binding
26	91.5	4.3	431	2	C86812 conserved hypothet
27	91.5	4.3	684	2	T33785 hypothetical prote
28	91.5	4.3	686	1	ALDYAT amyliase A (EC 3.2.
29	91.5	4.3	870	2	T01310 hypothetical prote

30	91.5	4.3	1416	2	T20823 hypothetical prote
31	90.5	4.2	244	2	T19750 hypothetical prote
32	90.5	4.2	403	1	Q0844E hypothetical prote
33	90.5	4.2	588	2	T37542 conserved hypothet
34	90.5	4.2	625	2	C75044 threonyl-tRNA synt
35	90.5	4.2	632	2	A82382 acetylhydroxy acid
36	90	4.2	1112	2	T47784 hypothetical prote
37	89.5	4.2	199	2	AD1230 phosphoglycerate m
38	89.5	4.2	376	2	B84277 hypothetical prote
39	89.5	4.2	397	2	T23642 hypothetical prote
40	89.5	4.2	407	2	A83279 hypothetical prote
41	89.5	4.2	416	2	S16658 mobilization prote
42	89.5	4.2	631	1	BHTLA hemocyanin chain a
43	89	4.2	217	2	H71507 hypothetical prote
44	89	4.2	333	2	C84792 hypothetical prote
45	89	4.2	335	2	B71485 probable tpr repea

ALIGNMENTS

RESULT 1
A56822
synaptonemal complex protein SC56 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 04-Mar-2000
C:Accession: A56822; S20742
R:Chen, Q.; Pearlman, R.E.; Moens, P.B.
Biochem. Cell Biol. 70, 1030-1038, 1992
A:Title: Isolation and characterization of a cDNA encoding a synaptonemal complex pr
A:Reference number: A56822; M0ID:93213429
A:Accession: A56822
A:status: preliminary
A:Molecule type: mRNA
A:Residues: 1-431 <CH>
A:Cross-references: EMBL:X65454; NID:957191; PIDN:CAA46449.1; PID:957192
A:Experimental source: testis
A:Note: sequence extracted from NCBI backbone (NCBI:P129004)
C:superfamily: rat synaptonemal complex protein SC56

Query Match	47.8%	Score 1019;	DB 2;	Length 431;
Best Local Similarity	53.6%	Pred. No. 1.5e-75;		
Matches	201;	Conservative	53;	Mismatches 91;
			Indels	30;
			Gaps	5;
QY	42	MPLESAVRAHLDKYSCHMAESKYLEISLRHLRLDSEAFCHRNCS--AAPPEPAAG	99	
DB	1	MPLATVAGHALDEQYEGESWRESARYLEAALRLHLRLDSEAFCHANCSPATSQPRAPG	60	
QY	100	LASVP-----ELRLFGGLLRRAHCLKRCCKGGLPAFRSQSPSREVLADPQRE	146	
DB	61	-----PDGDNDEGDEMDARELRLEFGHVLERRACLRCKRTTPAIVQVPPSQQLLDFQNR	116	
QY	147	PYKFLQFAVFRKANLLEKPAIAAHTFLKHPDDEMKKNMAVYYS-LPGADYIKDLETKS	205	
DB	117	PYGLYHAFHFKANRLKEKAAVAATFLQNRPKHETLAKVLYNVRMLDIGESLTDLEAQP	176	
QY	206	YESLFIPTAVRAVYNGENKRTISITDMELALDPFFKAFYCLACGSSRSKIKPKFYLISAD	265	
DB	177	YEAVFLQAVKLVNYSNGDFRSSTEHMERALADYMTVFACLCAGCEGAHGOVQFEKDYPAIAD	236	
QY	266	HYEVELECKTQCEENLETPVIGYVPEKFAVTMYHYLQFAVYKLLNDLKNAPCAVSYLLFD	325	
DB	237	LPEASLDCKKDCENLTPNNGFFVDFVATMTYHLYLQFAVYKLLNDVQAARSASTYLLFD	296	
QY	326	QNDKVMQNLVYVQYHRTWGLSDEHFQPRPEAVQFNVTTLOKELYDPAKENIMDDDEG	385	
DB	297	PKDSVMQNLVYVRFHRAKRLGEEDFQPREEAVLYNHQNSLELRDLFTHTMYLQSDSE-	355	
QY	386	EVEVEYVDLLELEET 400		
DB	356	-----MELEET 361		

RESULT 2
A64430
DNA-directed RNA polymerase (EC 2.7.7.6) subunit A' (intein-containing) - Methanococcus
N:Contains: Intein
C:Species: Methanococcus jannaschii
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 17-Mar-2000
C:Accession: A64430
R:Ball, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Ralch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kane, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1345 <BUL>
A:Cross-references: GB:067547; GB:L77117; MUD:g1591695; PID:g1591696; TIGR:MJ1042; PID:g
C:Genetics:
A:Map position: FOR974597-978634
A:Start codon: TTG
C:Superfamily: yeast DNA-directed RNA polymerase III 160K chain
C:Keywords: nucleotidyltransferase; protein splicing; transcription

Query Match 5.0%; Score 106; DB 2; Length 1345;
Best Local Similarity 23.8%; Pred. No. 3.3;
Matches 55; Conservative 34; Mismatches 82; Indels 60; Gaps 12;
QY 189 KSLGAEYITDLETKSYESLFIKAVRYNG-----EMKRISITDMELALDFFFAF 240
DB 714 KKMGIIPWIK-TAPKTIKKEFL--SAVFSESLTPKRNHGSFKSLSPKIKIEIF 769
QY 241 YECIACGSEGRKIDFDFYLSIDHYVELECKIOCEE-MLTPVIGSYPRKEFYATYH 299
DB 770 DE-----DFIKDKEMLKEFG-----TELKRVYEGSLRK--DCKTKVYASTYN 814
QY 300 YLQF---AYKKLDLNAACAVSYLF-----DONDKVMQNLVYYQYHRDTWGLSD 349
DB 815 HKEFFGRIGTYANKKETLARYAYEYLLTKEKYLKDRNIKKLENNTKFTFDK----- 867
QY 350 EHPQPREAVQFNVITLQKELYPAKENIMDDGEVVEYVDLLLEET 400
DB 868 -----FIKCKLKN--GFVKEKIVSIEETK-VDYVYDITITSET 903

RESULT 3
B95113
competence protein Coia [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: B95113
R:Nettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Reed, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95113
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-317 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75099.1; PID:g14972453; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0978

Query Match 4.9%; Score 104; DB 2; Length 317;
Best Local Similarity 22.1%; Pred. No. 0.7;

Matches 61; Conservative 38; Mismatches 119; Indels 58; Gaps 11;
QY 13 ALLCVACA----LRAGRAQYERYFSRFPDEMLPESAYRHALDKYSGEHWASGXYLE 68
DB 23 AYTCACGGQGLHLRQGSVPRTHFAKSLKDCDFFENSPHLANKESLYHMLKRETVQ 82
QY 69 ISLRHLRLRDSFAFCRN-----CSAAPQ---PEPAAGLAS--YPELRIFGLIRRAH 117
DB 83 LEYPLSELKQADYFVGNLALAEVQCSPLPKVLKERSEGRSGGYVLMIGOKLMLKE 142
QY 118 CLKRCQGLPAFROS-----QPSREVL-----ADFORBPYKFLQFAFRKAN--- 159
DB 143 RLRLQGSFLYFSQMGFYWELDKQVLLKLYITDYLKGLKLYQKESYGGSLLE 202
QY 160 --NLP-KAIAAHTFLKHPDEMKNMAYYKSL---PGAEDYIKDLETKSYESLPIR 212
DB 203 ILRLPYRKOKISHFTVSEDKDICRYIRQOLYYOVLFWMKEQEAAYOKG-----E 251
QY 213 AVRAYNGENMRTST-----DMELALPDEFKAFY 241
DB 252 NILTYGLKEWYQPIRPYVKGFEQIEDLTSTYYQHRY 287

RESULT 4
A97982
competence protein [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: A97982
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Buretti, S.; Dehoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balitz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97982
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-317 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99685.1; PID:g15458486; GSPDB:GN00174
C:Genetics:
A:Gene: coia

Query Match 4.9%; Score 104; DB 2; Length 317;
Best Local Similarity 22.1%; Pred. No. 0.7;
Matches 61; Conservative 38; Mismatches 119; Indels 58; Gaps 11;

QY 13 ALLCVACA----LRAGRAQYERYFSRFPDEMLPESAYRHALDKYSGEHWASGXYLE 68
DB 23 AYTCACGGQGLHLRQGSVPRTHFAKSLKDCDFFENSPHLANKESLYHMLKRETVQ 82
QY 69 ISLRHLRLRDSFAFCRN-----CSAAPQ---PEPAAGLAS--YPELRIFGLIRRAH 117
DB 83 LEYPLSELKQADYFVGNLALAEVQCSPLPKVLKERSEGRSGGYVLMIGOKLMLKE 142
QY 118 CLKRCQGLPAFROS-----QPSREVL-----ADFORBPYKFLQFAFRKAN--- 159
DB 143 RLRLQGSFLYFSQMGFYWELDKQVLLKLYITDYLKGLKLYQKESYGGSLLE 202
QY 160 --NLP-KAIAAHTFLKHPDEMKNMAYYKSL---PGAEDYIKDLETKSYESLPIR 212
DB 203 ILRLPYRKOKISHFTVSEDKDICRYIRQOLYYOVLFWMKEQEAAYOKG-----E 251
QY 213 AVRAYNGENMRTST-----DMELALPDEFKAFY 241
DB 252 NILTYGLKEWYQPIRPYVKGFEQIEDLTSTYYQHRY 287

RESULT 5
h70470
hypothetical protein aq_1989 - Aquifex aeolicus


```

QY 149 KELOFAVYKANKUNPKRAIAAATHELKHPDDEMKRNMAYKYSLGEAEYIDL----- 201
Db 217 KFINIDMEEFDELLETVAAFMESVAKFD---IKAGIOLQAYIPDSYTYLKLFPFSKER 272
QY 202 ETKSYSELFIAVRAYNGENNRSTITDMELALPDFEKAFA---YECIA--ACEGSEIREIK 254
Db 273 VLKGKRSIKIRFPVKGANMSEETIASORGWELPFPPYKKITDSDSNYMKMLDFLEBGD--- 328
QY 255 DEKOEYLSIAOHYV-EVELECKIOCEENTPYIGGSPVKKFATMYH--YLQFAY--YKLN 309
Db 329 NKYINIGISASHNIFELIYATRISE---VGA--IDSFFEMIEGMSLDCSTELSKMH 381
QY 310 DLKNAAP-C-----AVSYLLFPDQNDKVMQONLYYOYH---RDTWGLSDENHOPRPE 357
Db 382 DLILYTPVCDEAHFNNAIAYLYVRIDENTSEDNEMRFPFNKLKVGDKMNMIQELFLKSL 441
QY 358 AVQFENVYTLQKELYDFAKE-NIMDDDEGE 386
Db 442 GIKTLDNTTHKQ--DRNKESNIITSYEGK 469

```

RESULT 11
A71076
hypothetical protein PH0873 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: A71076
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki,
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: A71076
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-788 <KMW>
A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BAA29967.1; PID:d1030910; PID:g32572
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
C:Genetics:
A:Gene: PH0873

	Query Match	4.58;	Score 96;	DB 2,	Length 788;
	Best Local Similarity	24.44;	Pred. No. 11;		
	Matches 64;	Conservative 42;	Mismatches 98;	Indels 58;	Gaps
QY	153	FATRKANNLPKAIATAAHTFLIKHPDDEMKRKNAAVYYSLPGAEADYIKOLETKSYESLFIR	212		
Db	429	FIIYKAPEECCRWSPRIFFIVSRDMWDSKILLGLIYNPPIFGKKVKTLLD-----FIMQ	482		
QY	213	AVRAY--NGEN---WRISITFMELALDPDFEKAFECYLACEGSREIKDFDYSLADHY	267		
Db	483	AIEDYKNNGGANLAYYIIITLDEMILAHVEY--FADFISLVLESGBDEKGF--RRSIKLHD	538		
QY	268	VEVELECK--IGCEENLRP---VIGGYVEKEFVALMTHY----LOFAI---YKLDLKMAA	315		
Db	539	VDEVEEKGVYEKIPLPNLTIVGTWNIDE--FTTFMSPKVLRAVLIEFHEVDLENYP	955		
QY	316	PCAAVSYLEFDONDVVMQONLTVUYUHHNDTGMGSLDENHQPPEAAVOFPN-----VTTLQKE	370		
Db	596	PREEK---LDEDHVVALRNLILEDLRDGS-----KFLNTSKHNEINVAERE	637		
QY	371	LXDFAKENIMDDGEVVEVAYVD	392		
Db	638	L-DLSKVRLN---EVLEPER	654		

RESULT 12
E90139 conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus

C:Date: 24-Ma-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: E90139
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awaez, M.J.; C.
Jong, I.; Jeffries, A.C.; Kozeta, C.U.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redde-
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: E90139
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-240 <KUR>
A:Cross-references: GB:AE006641; NID:g1381332; PIDN:AAK40372.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS00007

	Query Match	4.5%	Score 95.5	DB 2	Length 240
	Best Local Similarity	23.8%	Pred. No. 2,4		
	Matches	54	Conservative	36	Mismatches 88; Indels 49; Gaps 10
QY	175	HPD----	DEMMRN----	MAYKSLPGADYIKDLETSYESLFRANRANGEN-----	221
DB	36	HDDGGLSIFLRLRNKWKIYYRS----	EDPIYEDPNISY-----	LIDANGEDIVGYI	85
QY	222	-----	WRTSITMEDALPPEFKA--FYECIACAGEGRSEIKDFQDFYLSIDHVEVEICKI		275
DB	86	HAKKGRKRLTGLSNAHPIYMRANITYDLFAINNGSVSRFSFDSNRPFDSTYLILEIRI			145
QY	276	OCENILTPYIGY PVEKEFVATMYHYLQFAYRYKLANDIKNAAPCAVSYLLFPDQNDKYQMOU			335
DB	146	LIESMNSP--FDYVYAI-----	TLDKLKDCSTLNSLSLI--YNNKEGSPIL		187
QY	336	VYVOYHRTDWG--LSDDEHQPRLPEAVQFENVNTYITKELCDPFAKNEM			380
DB	188	IGYVNRRLSKETNEEYIKLTYTDNKKRYLFSSITIKYLLGVADDEELIM			234

RESULT 13

H65011

probable transaminase (EC 2.6.1.-) b2379 [similarity] - Escherichia coli (strain K-1

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000

C:Accession: H65011

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.

.A.; Rose, D.J.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID: 97426617

A:Accession: H65011

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-412 <BIAT>

A:Cross-references: GB:AF000326; GB:U00096; NID:g1788718; PIDN:AMC75438.1; PID:g1788

A:Experimental source: strain K-12, substrain MG1655

A:Superfamily: aspartate transaminase

C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate

C:KeyWords: aminotransferase; phosphoprotein (lys) (covalently) status predicted

F:244/Binding site: pyridoxal phosphate

```

Query Match Similarity 4.5%; Score 95; DB 2; Length 412;
Best Local Similarity 19.8%; Pred. No. 5.4; Mismatches 126; Indels 132; Gaps 21;
Matches 77; Conservative 54;
OY 15 LCVACALNAGRAQVRYRS-FRSPRDELMPLESAVYRHAL-----DKYSGEHAESXGYLE 68
    | | | | | | | | | | | | | | | | | |
DB 58 LCTV---AQGRDTHGYSRSRIGPR-----LRAISRMYQDRYDVEIDPESEAIYT 104
    | | | | | | | | | | | | | | | | | |
OY 69 ISLR---LHRLRDSEAFCHRNCSAPOPEPAAGLASYPDELRLFGGLRAHCLRKCKOG 125
    | | | | | | | | | | | | | | | | | |
DB 105 IGSKGLHMLMATLID---HGDTLVLPN-----SYV-IHYGVAVIAGAVY-----RS 148
    | | | | | | | | | | | | | | | | | |
OY 126 LP-----AFKRSQPS-REVLADFORREPYKFLQAFYRKANNLPALAAATFE 171
    | | | | | | | | | | | | | | | | | |

```

Db 149 VPLVEGVDFNLELRAIRRESYPRKPMILIGFSPNPTAOCVELEFPE-----KVALAKRY 203
QY 172 LKHEDEMMKRMAY-----YKSLPGAEDYIKDELTSYSLSFIRAVRANG 219
Db 204 -----DVLVYHDLVIAITVDGKAPSIIMOVPGARDAVAEFFTIS-----KSTNM 248
QY 220 ENMRTSITDMELAPDF---FKAFYE-----CLACGSGRE-IKDFKDFYLSIAD 265
Db 249 AGWRIGFPMVGNKTLVSALARIKSYHDYGFPLQVAATAALEGDDQCVRDIAEQYKRRRD 308
QY 266 HYVEVLECKIOCEEMLTVIGYIPYKFAVATYAH-----IQPAYKILNDLK- 312
Db 309 VLVKGLH-----EAGMVMEMPRKASYVMAKIPEPYAANGSLEFAKKILNEAKV 356
QY 313 -----NAAPCAVSYLFDONDKVMQ 332
Db 357 CVSPGICRGDYGDTIHVRFALEIENRDIRQ 385

RESULT 14

S64361
hypothetical protein YGR066c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G4545
C:Species: Saccharomyces cerevisiae
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 29-Oct-1999
C:Accession: S64361
R:Medler, H.; Scharfe, M.; Medler, E.; Wambutt, R.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64356
A:Accession: S64361
A:Molecule type: DNA
A:Residues: 1-292 <MED>
A:Cross-references: EMBL:Z72851; NID:q1323087; PID:e243420; PID:q1323088; GSPDB:GN00007;
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YGR066c
A:Map position: 7R

Query Match 4.4%; Score 94.5; DB 2; Length 292;

Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 63; Conservative 29; Mismatches 87; Indels 93; Gaps 12;

QY 179 EMMKRMAYKSLPGAEDYIKDELTSYSLSFIRAVRANGENMRTSITDMELAPDFK 238
Db 14 ERAKRN-----DNIEIYTSAYVSL-----PSTDFR--LPHFKA 46
QY 239 AFYECIACGSGREI-KDFKDFYLSIADH-----YVEVLECKIOCE 278
Db 47 KEEDCDVYEECTNLVGKNAKYTSLGRHDLRPLRGSGSQSKYTYTV-EVKIDTV 105
QY 279 -----ENLTPVIGYIPYKFAVATYAHYLOFAVYKILNDLKNAAPC 317
Db 106 NLPLKDSRSLDPHTGFTTKNLTPVLDKY-VTLFEGYVINYQFPL-----C 153
QY 318 AVSYLFDONDKVMQON-----LVYYQYHRDTWGLSDEHF-QPRPAVOFFNVTTLQK 369
Db 154 SLHMPAEETLDPYMAQRESDCSHMKRFGHFGSDMWSLTERNFQYNHESAPEFMN---QR 209
QY 370 ELXDFAKENIMDDDEGEVVEYVDLLELETS 401
Db 210 YIYLKMKERFLDDEEDENIMLDDNHLLEGAS 241

RESULT 15

B81279

probable fibronectin/fibrinogen-binding protein Cj1349c [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: B81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barris

Nature 403, 665-668, 2000
A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni, reveals
A:Reference number: B81250; MUID:20150912
A:Accession: B81279
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-435 <PAR>
A:Cross-references: GB:AL139078; GB:AL11168; NID:96968723; PID:CA873776.1; PID:9696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1349c

Query Match 4.4%; Score 94; DB 2; Length 435;

Best Local Similarity 20.1%; Pred. No. 7;
Matches 59; Conservative 54; Mismatches 87; Indels 94; Gaps 14;

QY 135 SREVLADPQREPPYKFLQFAIFKAN---NLKRAIAAHT--FLKH---PDDEMKRMMA 186
Db 14 SRFKKIDFKRINDNILELSPDREFFIDLTRAMSAITAKFNKYNAPDFMLK--- 70
QY 187 YKSLPGAEDYIKDEL-----TKSYSLFIRAVRANGENMRTSITD----- 228
Db 71 YFSNA-----FLKEVKVLEGNRIILCFSVKANAKASYESKTYFEPTGANTVILTDKDL 125
QY 229 -----MELALPDFFAFYECIACGSGREIKDFKDFYLSIADHYVEVLECKIOQ 277
Db 126 IIEALRHIDKSYRVYKPNVILLESLSKPYKMDENFEIEIKDFADYF---SRKFTSYESKIKQ 182
QY 278 ENLTPVIGYIPYKFAVATYAHYLOFAVYKILNDLKNAAPCAVSYLFDONDKVMQONLTV 337
Db 183 IKNLKLA---QVDK-----KIONLOEL-----FSSLDEENSLKLKALEY 218
QY 338 YQYHNDTWGLSDEHQPPEA-VQFENVYTLQKELYDFAKENIMDDDEGEVVEY 390
Db 219 -----RKRADVLEFANLSYLK-----DIREFKLDDEFGKELEF 251

Search completed: June 8, 2002, 02:09:53

Job time: 6933 sec

SEA ID NO: 1
Database: N.GeneSeq-032802
AC NO: AAX00801

File Copy
Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2002, 21:01:30 ; Search time 562.42 Seconds
(Without alignments)
11817.104 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 ttccctctccctccctt.....ataaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.GeneSeq_032802.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT.*
25: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3870.2	100.0	3871	20	AAX60801
2	3870.2	100.0	3871	22	Human secreted pro
3	3870.2	100.0	3871	24	Human CDNA encodin
4	2051.2	53.0	2549	21	Human polynucleoti
5	1955.4	50.5	1961	22	Human prostate can
6	1474.4	38.1	1586	21	Human CDNA sequenc
7	695.6	18.0	818	22	Human CDNA cancer
8	466.4	12.0	618	21	Human CDNA clone (
9	406.4	10.5	408	21	Human colon cancer
					Human colon cancer

10	371.6	9.6	401	22	AAX50786	Human tumour assoc
11	369.4	9.5	400	16	AAX21281	Human gene signatu
12	355.6	9.2	407	22	AAX50808	Human tumour assoc
13	340	8.8	351	22	AAX11840	Human CDNA clone (
14	337.2	8.7	373	21	AAX43677	Human secreted exp
15	258.6	6.7	279	23	AAX58007	CDNA #683 encoding
16	239.2	6.2	6670	22	AAL02846	Human reproductive
17	239.2	6.2	7713	22	AAL02846	Human reproductive
18	238.4	6.2	8658	22	AAX69231	Human immune/haema
19	238	6.1	295	22	AAX82993	Human immune/haema
20	235.4	6.1	32351	21	AAX21307	Human low adenosin
21	235.4	6.1	32351	21	AAX35185	Human adenosine re
22	235.4	6.1	40298	21	AAX21311	Human low adenosin
23	235.4	6.1	40298	21	AAX35189	Human adenosine re
24	234.6	6.1	4275	22	AAX19408	Human nervous syst
25	234.6	6.1	4883	22	AAX79410	Human immune/haema
26	234.6	6.1	16997	22	AAX80033	Human immune/haema
27	234.6	6.1	16997	22	AAX86146	Human immune/haema
28	233.8	6.0	9620	22	AAL06207	Human reproductive
29	233.4	6.0	6138	22	AAX57456	Human liver cell s
30	233.2	6.0	763	22	AAX06411	Human CDNA clone (
31	232.6	6.0	2283	22	AAX14438	Human CDNA sequenc
32	232.6	6.0	5979	22	AAX78985	Human immune/haema
33	232.6	6.0	5979	22	AAX80951	Human immune/haema
34	232.2	6.0	5332	22	AAX71264	Human cervical can
35	232.2	6.0	567	22	AAX69839	Human polynucleoti
36	232	6.0	3775	22	AAX161061	Human immune/haema
37	231.4	6.0	1449	22	AAX86057	Human immune/haema
38	231.4	6.0	1450	22	AAX86062	Human immune/haema
39	230.8	6.0	2847	22	AAX18224	Human nervous syst
40	230.8	6.0	2847	22	AAX82073	Human immune/haema
41	230.6	6.0	2108	22	AAX76677	Human immune/haema
42	230.6	6.0	2108	22	AAX76677	Human immune/haema
43	230.2	5.9	646	22	AAX77815	Human immune/haema
44	230.2	5.9	11446	22	AAX90464	Human digestive sy
45	230	5.9	13559	22	AAX15144	Human nervous syst

ALIGNMENTS

RESULT 1	
ID AAX60801	standard; DNA; 3871 BP.
AC AAX60801:	
XX 09-AUG-1999	(first entry)
DE	Human secreted protein encoding DNA (clone bd306-7).
KW	Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW	nutritional activity; cytokine; cell proliferation; immune stimulation;
KW	hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW	anti-inflammatory; tumour invasion; ss.
OS	Homo sapiens.
XX	
PN	WO926961-A1.
PD	03-JUN-1999.
XX	
PF	24-NOV-1998; 98MO-US25149.
XX	
PR	23-NOV-1998; 98US-0197886.
XX	
PI	26-NOV-1997; 97US-0066804.
XX	
PA	(GEMT) GENETICS INST INC.
PI	Agostino MJ, Clark HF, Collins-Racie LA, Evans C;
PI	Reichel K, Jacobs K, Lavallee ER, McCoy JM, Metberg D;
PI	Steininger RJ, Treacy M, Wong GC;

THIS PAGE BLANK (0.3PT0)

Db 1681 ccaagcgctgcgaactcttgacttcagatgacatctgcctctgctccccaatgctc 1740
QY 1741 tgggattacagcgctgagccacatgcccggcctctcttcacacttaccctgtctct 1800
Db 1741 tgggattacagcgctgagccacatgcccggcctctcttcacacttaccctgtctct 1800
QY 1801 taccctacactctgttttaccacacttaccctgtctctctctctctctctctct 1860
Db 1801 taccctacactctgttttaccacacttaccctgtctctctctctctctctctct 1860
QY 1861 cccatgttcatacgtcctctcttaccatcttggttgaggagcagctctctcgtctg 1920
Db 1861 cccatgttcatacgtcctctcttaccatcttggttgaggagcagctctctcgtctg 1920
QY 1921 ttttttcttttcccaagaaatcagatattttttaaataagaaataccctctagaa 1980
Db 1921 ttttttcttttcccaagaaatcagatattttttaaataagaaataccctctagaa 1980
QY 1981 gatgataatgttgaanaacccctcttgctattgtcttccagattttagctctcttc 2040
Db 1981 gatgataatgttgaanaacccctcttgctattgtcttccagattttagctctcttc 2040
QY 2041 tcccacatccggagaaatggtggaagacataggttaatttccagcctccacaatgctc 2100
Db 2041 tcccacatccggagaaatggtggaagacataggttaatttccagcctccacaatgctc 2100
QY 2101 ttcaacttgcctgactgttaccatcttaccacacacacacacacacacacacacacac 2160
Db 2101 ttcaacttgcctgactgttaccatcttaccacacacacacacacacacacacacacac 2160
QY 2161 tctagaagtcagaaatgctgcttcttgcctccacatttgcattcccaatttcaacgctt 2220
Db 2161 tctagaagtcagaaatgctgcttcttgcctccacatttgcattcccaatttcaacgctt 2220
QY 2221 ccgattcccatcttgaagacatgtttagttataagcttcgagcttggaaaagaaagaaagc 2280
Db 2221 ccgattcccatcttgaagacatgtttagttataagcttcgagcttggaaaagaaagaaagc 2280
QY 2281 cagagctgtcctagtttcatcttcttccatgtaattatttagtaccatctgtgctc 2340
Db 2281 cagagctgtcctagtttcatcttcttccatgtaattatttagtaccatctgtgctc 2340
QY 2341 taagcatctaaccttggaactagagatactcacagaataacagggaaagttccctgtct 2400
Db 2341 taagcatctaaccttggaactagagatactcacagaataacagggaaagttccctgtct 2400
QY 2401 catgagcttaacttcttaacaggaagaaagataagccaatacataggaataatataac 2460
Db 2401 catgagcttaacttcttaacaggaagaaagataagccaatacataggaataatataac 2460
QY 2461 aaggtatcatctagttatctgctgtggaagaaataagaaagagggaggttaagaa 2520
Db 2461 aaggtatcatctagttatctgctgtggaagaaataagaaagagggaggttaagaa 2520
QY 2521 tccctgagagtgagctgcaagttttaaatggtggtcctcactgaggaatgtacgtttgaagcaga 2580
Db 2521 tccctgagagtgagctgcaagttttaaatggtggtcctcactgaggaatgtacgtttgaagcaga 2580
QY 2581 gacgttttaggagagtgatctctgagacaagagcatccagagcagagaaagaaagatgtgacactg 2640
Db 2581 gacgttttaggagagtgatctctgagacaagagcatccagagcagagaaagaaagatgtgacactg 2640
QY 2641 ccccaagtgagaaactgtctacgttgctcagagaaagcagagcagagcagagcagagcag 2700
Db 2641 ccccaagtgagaaactgtctacgttgctcagagaaagcagagcagagcagagcagagcag 2700
QY 2701 tgggacaggggtagaaatggaagagagagcggtcgggagagagcagagcgggtggaagcgcttgg 2760
Db 2701 tgggacaggggtagaaatggaagagagagcggtcgggagagagcagagcgggtggaagcgcttgg 2760
QY 2761 cttctgctaagtgagatggaagcactggaaggttgaagcagagagtgcttgaatg 2820

Db 2761 cttctgctaagtgagatggaagcactggaaggttgaagcagagagtgcttgaatg 2820
QY 2821 ttatatatttgcaaggggtcacttctagctgcaatattgtgaaataactttagtggaagagggc 2880
Db 2821 ttatatatttgcaaggggtcacttctagctgcaatattgtgaaataactttagtggaagagggc 2880
QY 2881 agaaagaaagagagagagcctgtttaggaagcactgcaaggttccagagcttggcctgggc 2940
Db 2881 agaaagaaagagagagagcctgtttaggaagcactgcaaggttccagagcttggcctgggc 2940
QY 2941 cacagacacagcaggttgcataatctagatatttttgaagaaagacaaatagatcttc 3000
Db 2941 cacagacacagcaggttgcataatctagatatttttgaagaaagacaaatagatcttc 3000
QY 3001 tgaagatttgaaatgtagagtgtaagaaagaaagagtgtaattgacatcaaggttttg 3060
Db 3001 tgaagatttgaaatgtagagtgtaagaaagaaagagtgtaattgacatcaaggttttg 3060
QY 3061 gctcgaatagcaggaagaaagatgtagcttaccagttacgttaataaggaagagtggtggt 3120
Db 3061 gctcgaatagcaggaagaaagatgtagcttaccagttacgttaataaggaagagtggtggt 3120
QY 3121 aagtaaggaattgtgtgcagaagcaggtgtctgtgttgtaagtggtgtcgtgcga 3180
Db 3121 aagtaaggaattgtgtgcagaagcaggtgtctgtgttgtaagtggtgtcgtgcga 3180
QY 3181 aatcaagtgagatctctcagagctcaggtctcagcagagcctcgaagacagagctcgaa 3240
Db 3181 aatcaagtgagatctctcagagctcaggtctcagcagagcctcgaagacagagctcgaa 3240
QY 3241 tgcacttgggtttaaattgttgggggtgctctcagaaaggaacacgttgaagaccttatacagtc 3300
Db 3241 tgcacttgggtttaaattgttgggggtgctctcagaaaggaacacgttgaagaccttatacagtc 3300
QY 3301 atttatgtgctgtgagaaagttctctgagagtggtgtgtgtacatttgaagcagaatgacttca 3360
Db 3301 atttatgtgctgtgagaaagttctctgagagtggtgtgtgtgtacatttgaagcagaatgacttca 3360
QY 3361 gttgaaggaagatctcttggaaagagagcgttgaagcatttgcagcctacacatctgtgttag 3420
Db 3361 gttgaaggaagatctcttggaaagagagcgttgaagcatttgcagcctacacatctgtgttag 3420
QY 3421 tctgttgggggttgggggtcctgagcactggtctgtgaaagagatcttgcagggcaccaca 3480
Db 3421 tctgttgggggttgggggtcctgagcactggtctgtgaaagagatcttgcagggcaccaca 3480
QY 3481 ggcgccctactgaaacatcagcatgtcagttgacatttaagccatgacgttgaaggggc 3540
Db 3481 ggcgccctactgaaacatcagcatgtcagttgacatttaagccatgacgttgaaggggc 3540
QY 3541 cactgagattgtctctgagttacttactgaaagaaacaaagaaagagccatgagtgagcc 3600
Db 3541 cactgagattgtctctgagttacttactgaaagaaacaaagaaagagccatgagtgagcc 3600
QY 3601 ttgggctctctggaagaaatggaatcagcgaagagcctgaagagaggttacccttaaggtc 3660
Db 3601 ttgggctctctggaagaaatggaatcagcgaagagcctgaagagaggttacccttaaggtc 3660
QY 3661 agagaanaacaaagagatgtgtgtcttctggaagcctgaggttcttattcaacactcattc 3720
Db 3661 agagaanaacaaagagatgtgtgtcttctggaagcctgaggttcttattcaacactcattc 3720
QY 3721 ccttctccaaataaagccacttgtgtagttgtggcctcctcgaaggttgaagcagaagagaga 3780
Db 3721 ccttctccaaataaagccacttgtgtagttgtggcctcctcgaaggttgaagcagaagagaga 3780
QY 3781 aagagcagcggttgggaaacaaagacatttctcgtcaatagccttggaaagaaagaaagga 3840
Db 3781 aagagcagcggttgggaaacaaagacatttctcgtcaatagccttggaaagaaagaaagga 3840
QY 3841 tagagtggttaaaaaataaaaaaa 3871
Db 3841 tagagtggttaaaaaataaaaaaa 3871

THIS PAGE BLANK (EXPTO)

THIS PAGE BLANK (1-3PTO)

PI Steininger RJ, Treacy M, Wong GG;
 XX
 DR WPI: 1999-357809/30.
 DR N-PSDB: AAX60801.
 XX
 PT New polynucleotides encoding secreted proteins
 XX
 XX Claim 10: Page 111-113; 133pp; English.

The invention relates to secreted proteins (AA17219-228) encoded by
 CC polynucleotides obtained from human fetal kidney, adult lung, adult
 CC kidney, adult brain, adult blood, adult testes, and fetal brain and
 CC murine adult bone marrow cDNA libraries. The secreted protein nucleic
 CC acid sequences (X6801-811) correspond to clones bd3067, g1283-6, 1,
 CC FK317-3, k213-2x, na316-1, n193-20, np164-1, pe204-1, ya1-1 and yb-1,
 CC (all clones are deposited as ATCC 98399). The PNs and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, hematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, anti-tumour inhibition activity. The PNs are
 CC also stated to be useful for gene therapy.

Sequence 401 AA:

Query Match 99.9%; Score 2128; DB 20; Length 401;
 Best Local Similarity 100.0%; Pred. No. 4.3e-184;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGRGGAALALALCVACALRGAQYERYSEFSPRDELMPLESAYRHADKYSGEHM 60
 DB 1 MEGYRGAALALALCVACALRGAQYERYSEFSPRDELMPLESAYRHADKYSGEHM 60
 QY 61 AEEGXGLESLRLHRLRSEAFCHNCNAPQEPAPAGLAPPELRGLIRAHCK 120
 DB 61 AEEGXGLESLRLHRLRSEAFCHNCNAPQEPAPAGLAPPELRGLIRAHCK 120
 QY 121 RCKOGLPAPRQSQPREVLADFORREPKYKLOFAYRKANLPAITAATFLIKHDDDM 180
 DB 121 RCKOGLPAPRQSQPREVLADFORREPKYKLOFAYRKANLPAITAATFLIKHDDDM 180
 QY 121 RCKOGLPAPRQSQPREVLADFORREPKYKLOFAYRKANLPAITAATFLIKHDDDM 180
 DB 121 RCKOGLPAPRQSQPREVLADFORREPKYKLOFAYRKANLPAITAATFLIKHDDDM 180
 QY 181 MRRNMAVYSLPGAEDYIKDLETKSYESLPIRAVAYNGENMPTSTIDMELLPDEFKAF 240
 DB 181 MRRNMAVYSLPGAEDYIKDLETKSYESLPIRAVAYNGENMPTSTIDMELLPDEFKAF 240
 QY 241 YECIAACBESREIKDFKDYLSIADHYVEVLECKIQCEENLPVIGGYEVEKFAVMYH 300
 DB 241 YECIAACBESREIKDFKDYLSIADHYVEVLECKIQCEENLPVIGGYEVEKFAVMYH 300
 QY 301 LQFAAYKLMDLNAPCAVSYLLPQONDKVMQONLVYQYHNDTWSLSHQPPEAVQ 360
 DB 301 LQFAAYKLMDLNAPCAVSYLLPQONDKVMQONLVYQYHNDTWSLSHQPPEAVQ 360
 QY 361 FENVTTLQKELYPAKENIMDDGEVEYVVDLLELETS 401
 DB 361 FENVTTLQKELYPAKENIMDDGEVEYVVDLLELETS 401

RESULT 2
 ID AAU38989 standard; Protein; 401 AA.

AC AAU38989;

DT 16-JAN-2002 (first entry)

man secreted protein bd306_7.

XX Human; secreted protein; antiinflammatory; immunosuppressive;
 KW neurotrophic; neuroprotective; antiarthritic; anticardiac; vulnary;
 KW cytostatic; antidiabetic; vitruce; antiferility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; anticancer; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement; vaccine.
 XX Homo sapiens.
 XX WO200175068-A2.
 XX 11-OCT-2001.
 XX 22-MAR-2001; 2001WO-US09369.
 XX 30-MAR-2000; 2000US-0539330.
 XX 04-DEC-2000; 2000US-0729674.
 XX (GENY) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie E, Collins-Racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 PI Clark H, Fehnel K, Werberg D;
 DR WPI: 2001-639363/73.
 DR N-PSDB: AAS59207.
 XX Secreted human proteins, useful as vaccine for treating various
 XX diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 XX nervous system disorders (e.g. stroke).
 XX Claim 2: Page 457-459; 619pp; English.
 XX The invention relates to novel human secreted proteins, the nucleic
 XX acids encoding them. The protein may exhibit cytokine, cell proliferation
 XX or cell differentiation activity or may induce production of other
 XX cytokines in certain cell populations and may exhibit immune stimulating,
 XX or immune suppressing activity, which is useful for the treatment of
 XX various immune deficiencies and disorders e.g. severe combined
 XX immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 XX systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 XX inflammation. The proteins are also useful in the treatment of diseases
 XX and disorders including tissue, skin and organ transplantation and in
 XX graft-versus-host diseases (GVHD). In the induction of tumour immunity,
 XX myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
 XX in the treatment of burns, incisions and ulcers; as well as in treatment
 XX of periodontal disease, osteoporosis or osteoarthritis, mediated by
 XX inflammatory processes, diseases of the peripheral nervous system,
 XX Alzheimer's, Parkinson's disease, Huntington's disease,
 XX amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 XX infarction of cardiac and central nervous system vessel e.g. stroke,
 XX sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 XX protein, having activin- or inhibin-related activities is useful as a
 XX contraceptive based on the ability of inhibins to decrease fertility in
 XX female mammals and decrease spermatogenesis in male mammals. The
 XX proteins and nucleic acids are also useful as food supplements. The
 XX present sequence represents a secreted protein of the invention.

Query Match 99.9%; Score 2128; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 4.3e-184;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (except)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2002, 20:54:15 ; Search time 3903.73 Seconds
(without alignments)
13383.793 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 ttcccttcctccctccctt.....ataaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hcc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	996.4	25.7	1079	9	AL569764
2	993.8	25.6	1029	9	AL545483
3	964.4	24.9	980	9	AL517454
4	954.4	24.7	1031	9	AL571267
5	936.2	24.2	964	9	AL546910
6	931.6	24.1	971	9	AL556976
7	924.4	23.9	1652	11	AK017797
8	918	23.7	971	9	AL572160
9	917.8	23.7	1005	9	AL514488
C 10	917.4	23.7	925	9	AL574681
11	896.6	23.2	1053	10	BM471592
12	889.6	23.0	916	9	AL541167
13	860	22.2	909	9	AL517455
14	858.8	22.2	973	9	AL543090
15	849.8	22.0	956	9	AL554750
16	816	21.1	916	9	AL550966
C 17	814.8	21.0	830	9	AL579089

C 18	804.8	20.8	863	9	AL580676
19	788	20.4	961	10	BI255097
20	777	20.1	827	10	BI760650
21	772.6	20.0	817	9	AL558186
22	771.4	19.9	774	9	AL598447
23	770.6	19.9	854	10	BI860855
24	765.2	19.8	846	10	BI253401
25	760.8	19.7	813	9	AL558966
26	759.6	19.6	933	10	BI858585
27	742.8	19.2	797	9	AU141715
28	741.6	19.2	1034	10	BM459357
29	737	19.0	846	10	BI260161
30	736.6	19.0	1088	10	BF025893
31	734.4	18.9	969	10	BG677743
32	731.8	18.9	1430	11	AK007624
33	731	18.9	782	10	BE877856
C 34	720.6	18.6	723	9	AL568770
35	715	18.5	846	10	BI862238
36	713.8	18.4	902	10	BF793387
37	710.6	18.4	934	10	BM453316
38	706	18.2	759	10	BI856439
39	705.2	18.2	707	10	BE747825
40	697.8	18.0	713	10	BE746070
41	695.8	18.0	760	10	BG825941
42	695.6	18.0	769	10	BG288789
43	695.6	18.0	818	9	AU132637
44	693	17.9	702	10	BI256839
45	684	17.7	762	10	BE744751

ALIGNMENTS

RESULT 1
AL569764/c 1079 bp mRNA linear EST 16-FEB-2001
LOCUS AL569764 LTI_NFL006_PL2 Homo sapiens cDNA clone CSDDI002Yp17 3
DEFINITION prime, mRNA sequence.

ACCESSION AL569764
VERSION AL569764.1 GI:12925427
KEYWORDS EST.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (Bases 1 to 1079)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

JOURNAL
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES
source
1. 1079
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDI002Yp17"
/clone_1ib="LTI_NFL006_PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact: Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 317 a 226 c 285 g 246 t 5 others
ORIGIN

Query Match 25.7%; Score 996.4; DB 9; Length 1079;
 Best Local Similarity 97.8%; Pred. No. 3.4e-142;
 Matches 1026; Conservative 4; Mismatches 16; Indels 3; Gaps 2;

QY 901 accacccacagttatagaagcgtatccggtttgagaattgtgtctaccatgtatcatt 960
 Db 1046 ACCTCACCCAGTATARK-AGCTATCCAGTGAAGAAAT-TGTGGCTACCAATGATCATTT
 QY 961 acttcagatttgcctatatataagttgaacagacctgaagaatgcacccccctgtgcagca 1020
 Db 989 AMTTGACAGTTTGCCATATATAAGTTGAGACAGACCTGTWAAATGACAGCCCCCTGTGACGCA 930
 QY 1021 gctatctgctcttttgatcagaatgaacaaagttcatagcagcagaacctgtgtattaccagt 1080
 Db 929 GCTATCTGCTCTTGTGATGAGATGACAGAGCTCATGACACAGACCTGCTGATTAACCACT 870
 QY 1081 acccaaggaacactttgggctctcgtgtagagcacttccagcccaagcactgaagactc 1140
 Db 869 ACCACAGGACACTTGGGGCTCTCGGATGAGCACTTCCACCCAGACCTGAGCAGCTTC 810
 QY 1141 agttcttaattgtaccacactccagaagaagctgtatgactttgtctaaagaaataata 1200
 Db 809 AGTTCTTAAATGATGACCACTCCAGAGAGAGCTGATGACTTGTCTAAGGAAATATATA 750
 QY 1201 tggatgatgtagaaggaagttgttgaataatgtgagatgacctcttggaaacttgagaaga 1260
 Db 749 TGGATGATGATGAGGAGAAATGTGAAATGTGATGAGACCTCTTGGAACTGAGAGAGA 690
 QY 1261 ccaactgacccacagcaacccaagaagacttctctctgtcgttcaggaacacagattctt 1320
 Db 689 CCACACTGACCCACAGCAACCAAGAGACTTCTCTTGGCTTCAAGAAACACAGATTTCTT 630
 QY 1321 tgtccttctcccaacagcccaagcgtgtgtatcactcagagcctctcttactctccaa 1380
 Db 629 TGTCTTTTCCCAACAGCCCGAGGCTGTGATACCTAAGAGCCTTCTTTACTCTCCAA 570
 QY 1381 gtgaagaagggaagcccgctctctcttaactgaatgtcatcaggggttgagcgtcttcc 1440
 Db 569 GTGAAAGGGAAGCCCGCTCTCTTCACTGATGATGATGATGATGATGATGATGATGAT 510
 QY 1441 atcttcaacactgacactcactgttcaacactcactcactcactcactcactcactcact 1500
 Db 509 ATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
 QY 1501 gtctcgtctcttgcacagcgtgagtgacatgagcagcttccacgtcactcactcactc 1560
 Db 449 GTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390
 QY 1561 gcccttgggttcaagaatctgtgtcatcagcctcccgagttaccgtgagttacagga 1620
 Db 389 GCGCTTGGGTTCAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330
 QY 1621 tttgccaacacagcccgactaattttgtatctttagtagagacggggttttgcatttgg 1680
 Db 329 TGTGCCACACAGCCCGGCTAATTTTGTATTTTGTATGAGAGAGGGGTTTTCCTGCTGCT 270
 QY 1681 ccaagcgtgtctcgaactcttgacttgaatgacatcactgtccttgcctccacagctgc 1740
 Db 269 CCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 210
 QY 1741 tgggatttaaggggttgagccacatgcgcgctcttcttcaacttcaacttcaacttct 1800
 Db 209 TGGGATTTACAGGGGTGAGCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
 QY 1801 tatcttcaacttcttcttcaacttcaacttcaacttcaacttcaacttcaacttcaacttct 1860
 Db 149 TATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 90
 QY 1861 cccatgttcaatagcgtccttcttcaacttcttgaaggcagctcctcctcgttg 1920
 Db 89 CCATGTTCATACCTGCTTCTTACCATTTTGGTTTGAAGGGCAGTCTTCTCTGCTGCTGCT 30

QY 1921 ttttttcttttcccaagaatacagat 1949
 Db 29 TTTTCTTATTTTNAATAAATAATATAT 1

RESULT 2

AL545483 1029 bp mRNA linear EST 16-FEB-2001
 LOCUS AL545483 LTI_NFL006.PL2 Homo sapiens cDNA clone CS01015YK22 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL545483
 VERSION AL545483.1 GI:12877964
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1029)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..1029
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS01015YK22"
 /lsize_type="LTI_NFL006.PL2"
 /note="Vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 218 a 320 c 285 g 204 t 2 others
 ORIGIN

Query Match 25.6%; Score 992.8; DB 9; Length 1029;
 Best Local Similarity 99.2%; Pred. No. 1.2e-141;
 Matches 1024; Conservative 4; Mismatches 1; Indels 3; Gaps 3;

QY 32 gtccctctctctctctctcttccgcccgcgagtgagccggggcgccggcgccgagc 91
 Db 1 GTCCCTTCT 60
 QY 92 gctgtcagcgtctgt 151
 Db 61 GCTGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118
 QY 152 cttaagcttccgagcgtcttcccaagggagcagatgataccgctcgtcagtcgagcgtcctcggca 211
 Db 119 CTACAGCTTCCGACGCTTCCACGAGGAGCACTGATATCCCTCGAGTGGCTTACCGGCA 178
 QY 212 cgcgctgacaagtagacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 271
 Db 179 CGCGCTGAGACAGTACAGCGGCGAGCACTGGGCCGGAAGCCTGGGCTTACCTGGAGATCAG 238
 QY 272 cctgtgagctgacacgcttctgtgcgagcagcagcagcagcagcagcagcagcagcagcagc 331
 Db 239 CTTGGGCTGACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 297
 QY 332 cgcgcgcagc 391
 Db 298 CGGCGCGACCGCGAGCCCG 357

Oy	392	ggagctgtgcgcccgcgcgaacttgcttaagcgccgcgaagaaggccttgcgaagcttcgcg	451
Db	358	gggcttgcctgcccccgcgcgacctgctcaaggcgtgcCAAGCAGAgggccttgcCAgCCTTCGg	417
Oy	452	ccaagtccccaccgaacgcgcgaagtgctgtgcggaactlccagcgccgcgcgaacccctacaagtt	511
Db	418	CCAGTCCACGCCCGACCGCGAGGTGCTGGCGGACTTTCCAGGGCCCGGAGCCTTACAAGTT	477
Oy	512	cctgcagttgcgttaacttccaagggcaataatctcccacaagccatcgccgtgtccaac	571
Db	478	CCTGCAGTTCCCTTACTTCCAAGGCAAAATATCTCCCCAAGGCAATCGCGCTGCATCACAC	537
Oy	572	ctttctactggaagcatctctgatgaagaaatgtgaaggaggaacatggaatatataaag	631
Db	538	CTTTTACTTGAAAGCATCTCTGATGATGAATGATGAGAAGGAACATAGGCAATTATTATAAGG	597
Oy	632	cctgcctgtgtccgaggaactacattaaagaccttgaaaccaagtcatalgaagcctgtt	691
Db	598	CCTGCCTGCTGGCGGAGCACTACATTAAAGACCTGAAAACCAAGTCATNTGAAGCCTGTT	657
Oy	692	catccgagcagttgcgggcataaacacytgtgagaactctgsgaacaatccatcacagacatga	751
Db	658	CATCCGACACAGTGCGGGCATACAACGGTGAGAACTGGAGAACATCATCATCAGACATGGA	717
Oy	752	gctgcgcccctccgagcctcttcaaagccttltagaagtgctcgcgaagccttggaaggttc	811
Db	718	GCTGGCCTTCCCGCACTTCTTCAAAGCCTTTTAGCAGTGTCTTCGCAAGCCTGGAGAGGTTTC	777
Oy	812	caggagagatacaaggaactcaagattcttaccttccatccatagacagatcatactgtaagat	871
Db	778	CAGGAGATATCAAGGACTTCAAGAGATTTTACTCTTCCATAGCAGATATTATGTGAAGT	837
Oy	872	tctggaatgcaaaaataacacgltgtgaagaaaacctcaacccccagttataaggaagctacccgt	931
Db	838	TCTGGAATGCAAAATACAGTGTGAAGAAACACTCAACCAGTTAATAGAGAGGCTATCCGAT	897
Oy	932	tgaagaatttgtgtgcacatgatactatacttaacttgagttgcctatltaaigtlaagca	991
Db	898	TGAGAAATTTGTGGCTACMATGTATCATTTACTTGTGAGTTTGCTATTATTAAATTTAACGA	957
Oy	992	cctgaagaatgcagccccctgtgtgaagtcagtcatactctgtccttttgatcsgaatgacaagt	1051
Db	958	CCTGAAGATCCAGGCCCTCTGTGCAGTACAGTATCTCTTTGTGATCAGRATGACAAGT	1017
Oy	1052	catgcagacagaa	1063
Db	1018	CATGCAGCAGAA	1029
RESULT	3		
LOCUS	AL517454/c		
DEFINITION	AL517454 LTI_NF1011_NBC1 Homo sapiens CDNA clone CS0DA003YFI1 3	980 bp	mRNA linear EST 13-FEB-2001
ACCESSION	AL517454		
VERSION	AL517454.1	GI:12780947	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Ll,M.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	Bp 191 91006 EVRY cedex - France		
FEATURES	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
source	Location/Qualifiers		
	1..980		
	/organism="Homo sapiens"		

```

/db_xref="taxon:9606"
/clone="CSODAD03XFL1"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notice="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(4d) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax: (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

```

Query Match	24.9%	Score 964.4	DB 9	Length 980
Best Local Similarity	99.4%	Pred. No. 2.6e-137		
Matches 974	Conservative	4	Mismatches 1	Indels 1
QY 945	gctacatgtaacatcttactctgacgtttgcgctatataaagtctgaacgacctgaagaatgca	1004		
Db 979	gctaccatgtaatcatkacttgcattgcatttatttaagtgaacgacctgaagaatgca	920		
QY 1005	gcccccgtgtcagtcacgctatctgctctttgatcagaatgacaaggtcagcagacaac	1064		
Db 919	gcccccctgtgacagtcacgctatctgctctttgatcagaatgacaaggtcagcagacaac	860		
QY 1065	ctgtgtgtatcaacagttaccacaagggagaaaccttggggccctctggaatgaacacttccagacc	1124		
Db 859	ctgtgtgtatcaacagttaccacaagggagaaaccttggggccctctggaatgaacacttccagacc	801		
QY 1125	agacctgaagcagttcagttcctttaaibgtgacacacctccacgaaggagctglatgacttt	1184		
Db 800	agacctgaagcagttcagttcctttaaibgtgacacacctccacgaaggagctglatgacttt	741		
QY 1185	gctaaggaataataatgatagtatgataaggaagaagttgttggaaatagtgtgatacctc	1244		
Db 740	gcttaaggaataataatgatagtatgataaggaagaagttgttggaaatagtgtgatacctc	681		
QY 1245	tttggaaactggggagaggaacagcttaagcccaacgaacaaagaagacttccctcttggcgttca	1304		
Db 680	tttggaaactggggagaggaacagcttaagcccaacgaacaaagaagacttccctcttggcgttca	621		
QY 1305	ggaacaacaagaattcttcttcttccccaacagccagagctgtgtatcactcaagacctt	1364		
Db 620	ggaacaacaagaattcttcttcttccccaacagccagagctgtgtatcactcaagacctt	561		
QY 1365	ctcttctactctccaaagtgtgaaaggggaagcccccgtctctctaactgtcatgtcatcaagg	1424		
Db 560	ctcttctactctccaaagtgtgaaaggggaagcccccgtctctctaactgtcatgtcatcaagg	501		
QY 1425	tgaagcctggcccttctcattcttcaacacgtcagccactcaatgttcaaacctatttctaac	1484		
Db 500	tgaagcctggcccttctcattcttcaacacgtcagccactcaatgttcaaacctatttctaac	441		
QY 1485	tcttttttgaatagagctcgcctctcttggcccaagcggcggagatgcataagcagcttctca	1544		
Db 440	tttttttttgaatagagctcgcctctcttggcccaagcggcggagatgcataagcagcttctca	381		
QY 1545	gtctactgtcaacactccgcctcttgggttcaaggaacttctgtcatcaagcctcccgagta	1604		
Db 380	gtctactgtcaacactccgcctcttgggttcaaggaacttctgtcatcaagcctcccgagta	321		
QY 1605	ctgtgtattacaagcagatgtgcaccacaagcccgctcaattttagttaaagcgg	1664		
Db 320	ctgtgtattacaagcagcctgtgcaccacaagcccgctcaattttagttaaagcgg	261		

```
QY 1665 ggttttgccatgttggccaggctgtctcgaaacttctgacttcagatgatccatctgct 1724
DB 260 GGTTCGCAATGTTGGCCAGGCTGTCTCGAACTTTCATTCAGATGATCATCTGCT 201
QY 1725 tggctcccaagctgtgagttacagcgctgagccacacagccgagctcttctcacc 1784
DB 200 TGGGCTCCCAAGAGTGTGGATBACAGCGCTGAGAGCCGATGCCGCTCTTTCTCAC 141
QY 1785 tttaacactgtctcttactcctcaacatctgttttcaacacttcaaccctgtcttctcat 1844
DB 140 TTTACACCTGTCTCTTCTATCCTCAATCTGTTTTCACACCTTCATCCCTGCTTCAT 81
QY 1845 gtttaacactgtctcccaagctcatagctgctcttcttaccatlttggtttaagagc 1904
DB 80 GTTACACCTGTCTCTTCCCAATTCATGCGCTTCTCTTCCATTTGTTGAAGGCG 21
QY 1905 agtctctctgtctgtttt 1924
DB 20 AGTCTTCTCTGCTTGT 1

RESULT 4
AL571267/c 1031 bp mRNA linear EST 16-FEB-2001
LOCUS AL571267 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1015YK22 3
DEFINITION AL571267 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1015YK22 3
ACCESSION AL571267
VERSION AL571267.1 GI:12928392
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1031)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1. 1031
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1015YK22"
/clone_1b="LTI_NFL006_P12"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 309 a 220 c 279 g 217 t 6 others
ORIGIN
Query Match 24.7%; Score 955.4; DB 9; Length 1031;
Best Local Similarity 98.7%; Pred. No. 6e-136;
Matches 1020; Conservative 5; Mismatches 2; Indels 6; Gaps 6;
```

```
QY 1026 ctgctctttgatcagaatgaaaggtacatcagcaagaacctggtgtattaccagtaccac 1085
DB 913 CTGCTCTTTGATCGAAGTACAGAGTCAATGACAGAGAAACCGGTGATATTACAGTACAC 854
QY 1086 agggacaacttggggccctctcgatgagcaacttccagcccaagcctgaagcagttc 1145
DB 853 AGGACACTTGGGGCCCTCTCGGATGAGCAGCTTCCAGCCCAACCTGAAAGCGTTGAGTTC 794
QY 1146 tttaatgacacacttccagaagagctgtatgacttgtctgaagaaatataatgagt 1205
DB 793 TTTAATGATGACACACATCCAGAGAGAGCTGATGACTTTCCTAAGAAATATATATGAT 734
QY 1206 gatgatgaggaagaagtgtggaatatagtgatagtacctcttgaaacgaggaagcaagc 1285
DB 733 GATGATGAGGAGGAAGTGTGGAATATGTGATGATACCTCTTGGAAATCGAGAGACACAGC 674
QY 1266 tagcccaagcaacaaagaagacttcccttggagcttgaagaaacaaagacttctggcc 1325
DB 673 TAGCCCAAGCAACCAAGAGACTTCTCTTGGGCTTCAAGAAACAGATTCCTTGTGTC 614
QY 1326 ttltcccaagcccaagctgttgtatacctcagaagccttcttacttccaaagtga 1385
DB 613 TTTTCCCAAGCCAGCCAGGCTGTGATACCTCAGAGCCTTCTTACTCTCCAAAGTGA 554
QY 1386 agggaaagcccgctctcttaactgtcatgtcatcaagggtgagcctgtcttctatctt 1445
DB 553 AGGGAAGCCCGCTCTCTCTTAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 494
QY 1446 cacacccgcaacatcagttatcacaacctatcttctaac -tttttttgaatgagatct 1504
DB 493 CACACCTGCAACCTCATATGTTTACACCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 434
QY 1505 cgcctctctgcccagcgtgga -gtgcaatgagcaagcttctcagcactgaacctcgcgc 1563
DB 433 CGCTCTGTGGCCAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 374
QY 1564 tcttggttcaagaatctgtcatcatcagccttccagagtaactgtggaattacaggaatgt 1623
DB 373 TCTTGGGTTCAAGCAATCTCTGTCATCAGCCTCCGAGTACCTGAGGATDCAAGCAATGT 314
QY 1624 gccacacgcccagctaatcttgaattttgtagtagaagaggttttggcatttggcca 1683
DB 313 GCCACCAAGCCCGGCTAATTTTGTATTTTGTATGATGATGATGATGATGATGATGATGATGAT 254
QY 1684 ggcgtgctcgaactctgtacttgaatgatactcactgtgaccttggccctcccaagctgctg 1743
DB 253 GGCTGTGCTGAACACTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 194
QY 1744 gattacagcgctgagcaacatgcccgccctcttctcaaccttcaacctgtctctat 1803
DB 193 GATGACAGGCGTGAGCCACATGCGCCGCGCTTCTTCACTTCACTTCACTTCACTTCACTTCACT 134
QY 1804 cctcaactgttttcaacacttcaaccttcaaccttcaaccttcaaccttcaaccttcaacct 1862
DB 133 CCTACACATGCTTTTCAACACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 74
QY 1863 catgttaacagctgcttcttcaaccttgaattgttgaagggcagcttctctgtctgt 1922
DB 73 CATGTTCATAGCTGCTCTTCTTACCATTTGTTGATGAAGGCACTCTTCTGCTTGT 14
QY 1923 tttttgttttcc 1935
DB 13 TTTTGTGTTTCC 1

RESULT 5
AL546910 964 bp mRNA linear EST 16-FEB-2001
LOCUS AL546910 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1032YH24 5
DEFINITION AL546910 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1032YH24 5
ACCESSION AL546910
VERSION AL546910.1 GI:12880487
```


[illegible]

DB	579	ACACCTATCTTGTGAAGCATCCAGATGACGAGATGATGAAGAAACATGGAGTATTTATTA	638
QY	628	aaagccctgcttgctgcgaagagctatcatlaaagacctgaaacccaagtcatactgaagcc	687
DB	639	AGAGCTGTGCTTGAGACCCGAGGACCACTTTAAAGACTTGGAAACCAAGTCGTACGAGAAC	698
QY	688	tgctcatcccgagcagctgcggagcatacaacogtfgagaaactggaagaacatcatcacagaca	747
DB	699	TGTTTGTCCCTGCTGGCTGGGGGCTTCAACGCGGGAGACGTGGGAACGTCATTTCCGACA	758
QY	748	tggagcttgagcccttcgagctctccaaagcctttaaagagtgctgcagccttcgag	807
DB	759	TGGAGCTGCGCTTCCCGACTTCTCCAAAGCCTTTACGAGAGTGCTTGCTGCTTCGAGG	818
QY	808	gtccagaggagatacaaggagcttcaaggatlttctaccttccataagacaagatcatlaag	867
DB	819	GGGTGGCGGAGATCAAGACATTCAGAGACTTGTACTGTCATACAGATCCTATGTTGG	878
QY	866	aagcttcgtgaatgacaaaataacagtgctgaagagaacctaacccagttataggagctatc	927
DB	879	AAGTTCTGTGAGTGAAGATTGCTGTGAGAGAACCTTCAACCCAGTCAATGAGAGCTATC	938
QY	926	cggttgagaagaatttgctgctaccatgctatcatcttacttgcaagttgacctataagttga	987
DB	939	CCGTGGAGAAATTTGTGGCGACCATGTACACTATTTCAGTTTCGTTATTCAGAGTTTGA	998
QY	988	acgagactgaagaatgacagcccccctgtgagtcagctatctgctcttctatcagaatgaca	1047
DB	999	ANGATCTGAGAGATTCAGACGCCGCTGTGCCGTGACTGCTTCTTTACCAAGATGACA	1058
QY	1048	aggtcatgcaagcagaacctggtgctatctaacagctaccacaggaagaaccttggggcctcgg	1107
DB	1059	GGGTATGATGAACAGAACCTGGTGTACTGTACGTATACACCGGGACAAGTGGGGCTCTCGG	1118
QY	1108	atgagcacttccagcccgagacctggaagcagttcaagttctttaaagtgcacacatccaga	1167
DB	1119	ANGACACTCTCCAGCCCAAGCCGGAAGAGTTCACATTTTAAATGTGACGAGCTCCACA	1178
QY	1168	aagagcctatgactcttgctaaagaaaataatgatatgatatgagatgaggaagaagttgtg	1227
DB	1179	AGGAACCTGTACGACTTCCGTCTCAGAAACACCTAATGTGATGACGATGAGGAGAGGTGTGG	1238
QY	1228	aatactgagatgacctcttcttgaaactgtgagaagaccagctagccacagaaccaaagaga	1287
DB	1239	AGTATGTGGACGACTGTTGTGAGACGGAAGAGTGTGCTAGTCACACAGGCTTAAGGAAC	1298
QY	1288	cttccct 1293	
DB	1299	CTCTCT 1304	
RESULT 8			
AL572160/c			
LOCUS	AL572160	971 bp	mRNA
DEFINITION	AL572160 LTI.NFL006.P12 Homo sapiens CDNA clone CS0DI03YH24 3		linear EST 16-FEB-2001
ACCESSION	AL572160		prime, mRNA sequence.
VERSION	AL572160.1	GI:12930160	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequenage		
	BP 191 91006 Evry cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
FEATURES			

[illegible]

REFERENCE	1 (bases 1 to 925)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayres, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES	
source	Location/Qualifiers 1..925 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0D1065YP05" /clone_id="L11_NFL006_PL2" /tissue_type="placenta" /note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	275 a 198 c 251 g 193 t 8 others
ORIGIN	
Query Match	23.7%; Score 917.4; DB 9; Length 925;
Best Local Similarity	98.9%; Pred. No. 3.7e-130;
Matches 915; Conservative	7; Mismatches 3; Indels 0; Gaps 0;
QY 1029	ctctttgatgagaatgaagaagtcacgtcagcagaacacctggtgtattaccagttacacagcag 1088
Db 925	ctctttgatcgaatgaagaagtcacgtcagcagaacacctggtgtattaccagttacacagcag 866
QY 1089	gacactttgggacctctcgtatgtagaaccttccagcccaagacctgaagcagttcagttcctt 1148
Db 865	gacactttgggggacctctcgtatgtagaaccttccagcccaagacctgaagcagttcagttcctt 806
QY 1149	aatgtgaccacaacctccagaagaagcgtatgatgacttgcctaagaataataatgatgatg 1208
Db 805	aatgtgaccacacacttccagaagaagcgtatgatgacttgcctaagaataataatgatgatg 746
QY 1209	gatgaggggaagaagtgttgaataatgtgatagccctcttggaaacttggaggaagcagcag 1268
Db 745	gatgaggggaagaagtggtggaatgtgatgacaccttcttggaaactgagagacacagctag 686
QY 1269	cccaacagcaacaaagaagacttccctctcttggcggttcagaagaacacagattccttgcctt 1328
Db 685	cccaacagcaacaaagaagacttccctctcttggcggttcagaagaacacagattccttgcctt 626
QY 1329	tcccaacagcccaagcgtgttgataccacagagccttctcttactctccaagaatgaag 1388
Db 625	tcccaacagcccaagcgtgttgataccacagagccttctcttactctccaagaatgaag 566
QY 1389	gaagcccccgctctcttaactgcatagtcatacaggggtgagcttgctcttccattctaac 1448
Db 565	gaagcccccgctctcttaactgcatagtcatacaggggtgagcttgctcttccattctaac 506
QY 1449	acctgccaacctcaltgtcaaacctatctcttccaccttctttagatgtagtctgct 1508
Db 505	acctgccaacctcaltgtcaaacctatctcttccaccttctttagatgtagtctgct 446
QY 1509	ctcttgcacagcgtgagtgcaatgcaagcttctcagctcaactgcagacctccgctcttg 1568
Db 445	ctcttgcacacacgtgagtgcaatgcaagcttctcagctcaactgcagacctccgctcttg 386
QY 1569	ggttcaagcaatttcgctgcatacagccctccgagactcttgatatacaggaatgagcag 1628
Db 385	ggttcaaaccaatttcgctgcatacagccctccgagactcttgatatacaggaatgagcag 326

Oy 1629 cagcccgagcaatttctgattttagtagaagaggggtttgcacatgttgcagagctg 1688
Db 325 CACGCCCGGCTAATTTTGTATTTTGTAGAGAACGCGGGGTTTGGCATGTTGGCCAGGCTG 266
Oy 1669 gctcgaactcttgactcagaatgacatcctgcttgcccccacagtgctgagatta 1748
Db 265 GCTCGAACTCTTGACTTGAGATGATCCATCTGCTTGCCCTGCCACAGTGTGGAGATGA 206
Oy 1749 cagcgctgagccacacatcccgccgctcttctcacccttaacacctgtctcttaaccctca 1808
Db 205 CAGCGGKRGSGSSATGCGCGGCTCTTCTCACCCTTTACACCTGTCTTTATATCCCA 146
Oy 1809 catctgtttcacccttcacatccctgtctctcctcaatgttcaacatgtctctcccatgtt 1868
Db 145 CATCTGTTTTCACACCTTCATCCCTGTCTCTCATGTTCACATGTGTCTTCCCATGTT 86
Oy 1869 catagctccttcttaccatttggtttgaaggacacatctctctgtctgttttttg 1928
Db 85 CATACCTGCTTTCTTTACATTTTGTTTGAAAGGCGAGCTTCTGCTGTGTTTTTTTG 26
Oy 1929 ttttccagaataatcagtatatt 1953
Db 25 TTTTCCAGAAATCAGATATATT 1

RESULT 11
BM471592 1053 bp mRNA linear EST 05-FEB-2002

LOCUS AGENCOURT.646454 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5581397
DEFINITION 5', mRNA sequence.

ACCESSION BM471592.1 GI:18520634
VERSION BM471592.1
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1 (bases 1 to 1053)
NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

COMMENT

Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: LMNL2341 row: d column: 06
High quality sequence stop: 752.

FEATURES

Location/Qualifiers
1..1053

SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5581397"
/clone_id="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPORE6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 244 a 279 c 245 g 278 t 7 others
ORIGIN

Query Match 23.2%; Score 896.6; DB 10; Length 1053;
Best Local Similarity 96.9%; Pred. No. 5.3e-127;
Matches 965; Conservative 0; Mismatches 24; Indels 7; Gaps 5;

Oy 747 atgagagtcgcccctccgactcttcaagaccttaagagtgctctgcagcagctgcag 806

Db 1 ANGGAGCTGGCCCTTCCGACTTCTTCAAAAGCTTTTACGATGTCTGACACCTGCCAG 60
Oy 807 gttcccaaggagatcaagacttcaagattctcacttccatagagatcatatgta 866
Db 61 GGTCCAGGAGATATAGGACTTCAAGATTCTCTACTTTCCATAGCATATATATGTA 120
Oy 867 gaagttctgaaatgcaaaataacagtgatgaagaaaccccaacccagttataggagctat 926
Db 121 GAAGTTTTCGGAATGCAAAATACAGTGTGAAGAACTACACCCAGTTATAGAGAGCTAT 180
Oy 927 ccggttgagaaatttgygtacacatgatacttctgcaagtttgctatataagctg 986
Db 181 CCGGTTGAGAAATTTGTGGCTAACATGTATATTAATTGCAAGTTTGGCTATATTAAGTTG 240
Oy 987 aacgacctgaagaatgacgccccctgtgacgtcaacatactcgtctcttgatcaagatgac 1046
Db 241 AAGGACCTGGAAGATATGACAGCCCTGTGTGACATCTATCTGCTTTGATCAGAAATGAC 300
Oy 1047 aagttcatgacagcaaacctgtgtatatacagtaaccaaggaacatttggtccctcgc 1106
Db 301 AAGTCTATGACAGACAGAACTGGTATTAACAGTACACAGGAGACGTGGGGCTCTCA 360
Oy 1107 gatgagcacttccagccccagacctgaagcagttccttaatgtacacactccag 1166
Db 361 GATGAGCACTTCCAGCCAGACCTGAACTGACATGATGATTAATGTACACACTCCAG 420
Oy 1167 aagagctgatatgacttctgcaagaagaataataatgatatgatatgaggaatgtg 1226
Db 421 AAGGAGCTGTATGACTTTTGCTATAGAAATATATATGATGATGATGAGGAGAAATGTG 480
Oy 1227 gaatatgtgatgacaccttctgaacttgagagagacagctagcccaacgaacaaagag 1286
Db 481 GAATATGTGATGATGACCTTTGGAACTGAGGAGACACTGACCCACAGCAACAAAGAG 540
Oy 1287 acttctcttggtggttcaaggaaacaaatcttctgtctcttcccaacacccaggtg 1346
Db 541 ACTTCTCTTGCGCTTCAAGAAACACAGATCTTGTCTTTCCCAACAGCCAGGCTG 600
Oy 1347 ttgatacctcagaagcctctctcttctcctcaaaagtgaaggaagcccgctctcta 1406
Db 601 TTGATACCTCAGAGGCTTCTCTTACTCTCCAAATGAAAGGAAGCCCGCTCTCTTA 660
Oy 1407 actgcatgtcatcagggttgagcctgctcttctctacttctcaacctgcacactatgttc 1466
Db 661 ACTGATGTATACAGGAGGTGAGCCGCTTCTCTTCAACACTGACACTCATGTTC 720
Oy 1467 aaactatcttcttactcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1526
Db 721 ACACCTATCTTCTCACCCTTTTGTGAGATGAGATCTCGCTCTTCCAGGCTGAG 780
Oy 1527 tgcgaatgcaagttctcagctcactgcaacctcc-gcctcttggttcaagcaattctgc 1585
Db 781 TGCATGCGCAGCTTCTCAGCTCAGCTCAGCAACTCCAGCCTCTTGTGGTTCAGCAATCTGC 840
Oy 1586 tgcatacagcctcccgagtaacctg-gaatcaaggaatgtgcacacag-cccagtaact 1643
Db 841 TGCATCAGCCCTCCGAGTACCTGAGNATTTACAGGATGTGCACACAGCCCGGCTAANT 900
Oy 1644 ttgtatctttagtagaga---cgagggttgcgaatgttgcagcagctgtgtctgaacctt 1700
Db 901 TTGATTTTAAANMANAAGACGGGGTTTGCATGTTTGGCAGAGCTGTCTCGAATCT 960
Oy 1701 t-gactcatgatatcatctgccttgccctccac 1735
Db 961 TGGACTTCAGATGATCATCTGCTTGGCTTGGCCCTCC 996

RESULT 12

AL541167 916 bp mRNA linear EST 16-FEB-2001
LOCUS AL541167 LTI_FL002_P11 Homo sapiens cDNA clone CS0DE005F12 5 prime
DEFINITION , mRNA sequence.

ACCESSION AL5411167
 VERSION AL5411167.1 GI:12871971
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 916)
 AUTHORS Li,W.B., Gruber,C., Jesssee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..916
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DE005F12"
 /clone_id="L71_F1002_P11"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMWSPORT 6; 1st strand
 cDNA was primed with a NotI-cdo(cdy) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMWSPORT 6 vector. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifestech.com URL :
 http://fulllength.invitrogen.com"

Query Match	23.0%:	Score	889.6:	DB.9:	Length	916:
Best Local Similarity	99.2%:	Pred. No.	6.4e-126:			
Matches	911:	Conservative	3:	Mismatches	2:	Gaps
QY	65	ggaagccgggggcccgggggcccggcgcgctctatagcctctgtctgtatgctgcgtgcgcct	124			
Db	1	GGAGCGGGGGGGCCCGGGGGCCCGCGCCTGTCTAGCGCTGCTGTCTGCGCTTGGCGCCT	60			
QY	125	gcgcgcgcggcgcgccgccaatacgaacgctatagctcttcgcgcagctctccacgggaagact	184			
Db	61	GGCGCGCGGGGCGC-CCCAATACGAACGGCTACAGGCTCCGAGCTTCCACGGGAAGACT	119			
QY	185	gatgcgcgtctgagtgcgcctaccggcagcgctgtgaagaagtaacagggcgagactgggc	244			
Db	120	GATGCGCGCTCAAGTCTGGCGCTTACCGGCACGCGCTGTGCACAAATACAGCGCGCGAGACTGGCG	179			
QY	245	cgaagagcctkagctacacctggagatacagctctcgagtgacacgctgtgcgcgcgaacgga	304			
Db	180	CGAAGCGCTGGGCTACCTCTGGAGATCAGCTCGCGGGCTGGACCGCTTGCTGCGCGAGAGC-A	238			
QY	305	ggcctctctgcacacggccaactgcagcgccggcgccgagcccgagcccgccgcgcgcgtgc	364			
Db	239	GGCCTTCTGCGCACCGCACTGCAGGCGCGCGCGAGCGCAGCGCGCGCGCGCGCTGCG	298			
QY	365	cagctatcccgagctgtgcctctctcgaggcgctgtctgcgcgcgcgcgcgcgcgcgcgcgc	424			
Db	299	CAGCTATCTCCAGAGCTGCCTCTCTTCGGGGGCGCTGCTGCGCGCGCACTGCTCTCAACGG	358			
QY	425	ctgcaagcaggggacctgcagagctctcgccaatccacgcccagccggcgagtgctcgcgga	484			
Db	359	CTGCAAGCAGG6CCCTGCGAGCTTTCGGCAATCCCAAGCCCAAGCGCGAGGCTGCGCGGA	418			
QY	485	cttccagcgccgcgagccctacaagtctctgagcttcttacttcaaggaataatcct	544			
Db	419	CTTCCAGCGCGCGAGCGCCTTCAAGTTCTCTCAGTTCGCTTACTTCAAGGCAATAATCT	478			
QY	545	ccccaaagcatcgcgctgtctcaacaaccttctatgaagcatccctgatagaagaatgat	604			

Db	479	CCCCAAGCCATGCGCGCTGCTCACACCTTTCTTACTGAAGCATCCTGATGATGAATGAT	538
Qy	605	gaagagagacatgycataattaaagagccctgctgctgctgcgagagacatcaataaagacct	664
Db	539	GAAGAGGAACATGCGATTTATTAAGAGCCTGCGCTGTGTCGCCAGAGCATATTAAAGACT	598
Qy	665	ggaaaaccaagtcataatgaagccctgctccagcagctgcgagacatacaacgctgagaa	724
Db	599	GGAAACCAAGTCATATGAAAGCCTGTTTCATCCGACGACATGGCGGCGATCAACACGGTAGAA	658
Qy	725	ctggaggaacatccatcaacaagacatggaactgagccctcccgactcttcaagccttta	784
Db	659	CTGGGAACATCTCCATCACAGACATGAGACTGCGCCCTTCCGCACTTCTTCAAGCCCTTTAA	718
Qy	785	cgaagctctcgagagctcgagaggtctcgaagagaaatcaagaagcttcaagatcttact	844
Db	719	CGAGGTCTCGAGCGCTCGAGGGTTCACAGGAGATCAAGGCTTCAAGGATTTTACCT	778
Qy	845	tcccatagcagatcatatgtagaagctctggaatgcaaaatcacatagcttgaagaagacct	904
Db	779	TTCCATTACAGATCATTATTGTAGAAGTTCTGGAATGCAACATACAGTGTGAAGAAGACT	838
Qy	905	caaccacagttataggaagcgtatccggttgagaatcttgctacacatgatalcatlact	964
Db	839	CACCCCACTTATAGAGAGCTATCCCGTGTAGAAATTTGTGGCTACCATGATCATCTACTT	898
Qy	965	gcacgttgcctatatata 982	
Db	899	GCAGTTGCCCTATTATTA 916	

RESULT	13
LOCUS	AL517455
DEFINITION	AL517455 LTL.NF1011.NBC1 Homo sapiens cDNA clone CS0DA003YF11 5
ACCESSION	AL517455
VERSION	AL517455.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES	Location/Qualifiers
Source	1..909
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="CS0DA003YF11"
	/clone_id="LTL.NF1011.NBC1"
	/sex="male"
	/issue_type="neuroblastoma cells"
	/lab_host="DHL0B"
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiangellifotech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	195 a 283 c 251 g 176 t 4 others

Sat Jun 8 10:13:32 2002

us-09-729-674-1.rst

Page 14

Search completed: June 7, 2002, 22:07:46
Job time: 4411 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2002, 20:58:20 ; Search time 6829.55 Seconds
(without alignments)
11861.195 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 ttccctctccctccctt.....ataaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	2237	57.8	2307	9	HSJ6470	AJ006470 Homo sapi
2	1955.4	50.5	1961	9	AK001634	AK001634 Homo sapi
3	1370.4	35.4	1439	9	BC008745	BC008745 Homo sapi
4	944	24.4	1685	10	MMAJ6469	AJ006469 Mus muscu
5	497.6	12.9	1665	5	GGCARR	X97607 G.gallus mr
6	467	12.1	1889	9	BC001047	BC001047 Homo sapi
7	467	12.1	2037	9	BC007942	BC007942 Homo sapi
8	467	12.1	2443	9	AK056085	AK056085 Homo sapi
9	466.6	12.1	1745	9	BC011701	BC011701 Homo sapi
10	465.8	12.0	2347	6	AX335490	AX335490 Sequence
11	465.8	12.0	2347	9	HSU47621	U47621 Homo sapien
12	465.4	12.0	2079	9	HSJ250583	AJ250583 Homo sapi
13	403	10.4	1407	10	RNSC65MR	X65454 R.norvegicu
14	369.4	9.5	400	11	G27784	G27784 human STS S
15	355.6	9.2	407	6	AX150162	AX150162 Sequence
16	313.6	8.1	330	6	AX150140	AX150140 Sequence
17	263.4	6.8	274	6	AX333823	AX333823 Sequence
18	258.6	6.7	279	6	AX261032	AX261032 Sequence
19	252.2	6.5	157336	9	AC008810	AC008810 Homo sapi
20	250.6	6.5	257967	2	AL365337	AL365337 Mus muscu
21	249.6	6.4	164217	9	AC093264	AC093264 Homo sapi
22	249	6.4	200240	2	AC011618	AC011618 Homo sapi
23	247.8	6.4	157375	9	AC025207	AC025207 Homo sapi
24	247.8	6.4	178420	9	AP003548	AP003548 Homo sapi
25	247.8	6.4	199569	9	AC016250	AC016250 Homo sapi
26	247.8	6.4	211076	2	AC012214	AC012214 Homo sapi
27	247.4	6.4	167330	2	AC021822	AC021822 Homo sapi
28	247.4	6.4	192361	2	AC025678	AC025678 Homo sapi
29	246.6	6.4	190615	2	AL390994	AL390994 Homo sapi
30	245.8	6.3	169089	9	AC008751	AC008751 Homo sapi
31	245.8	6.3	176141	2	AC024393	AC024393 Homo sapi
32	245.6	6.3	192441	2	AC073404	AC073404 Homo sapi
33	245.2	6.3	164758	2	AC024393	AC024393 Homo sapi
34	245.2	6.3	187086	2	AC104982	AC104982 Homo sapi
35	245.2	6.3	188659	2	AC022169	AC022169 Homo sapi
36	245	6.3	76141	2	AC024393	AC024393 Homo sapi
37	243.8	6.3	178328	9	AC092547	AC092547 Homo sapi
38	243.6	6.3	159597	9	AC008747	AC008747 Homo sapi
39	243.6	6.3	159738	2	AC058855	AC058855 Homo sapi
40	243.6	6.3	165841	2	AC011276	AC011276 Homo sapi
41	243.4	6.3	170805	9	AL133321	AL133321 Human DNA
42	243.4	6.3	196720	9	AP002026	AP002026 Homo sapi
43	242.8	6.3	145299	2	AF228728	AF228728 Homo sapi
44	242.8	6.3	169740	2	AF186193	AF186193 Homo sapi
45	242.8	6.3	175256	2	AC090775	AC090775 Homo sapi

ALIGNMENTS

RESULT 1
HSJ6470
LOCUS HSJ6470 2307 bp mRNA linear PRI 23-MAR-2000
DEFINITION Homo sapiens mRNA for cartilage-associated protein (CASP).
ACCESSION AJ006470
VERSION AJ006470.1 GI:3687321
KEYWORDS cartilage-associated protein; CASP.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 2307)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Tonacchini,L., Morello,R., Montlicone,M., Skaug,J., Scherer,S.W.,
Cancedda,R. and Castagnola,P.
TITLE cDNA cloning, characterization and chromosome mapping of the gene
JOURNAL encoding human cartilage associated protein (CPRAP)
MEDLINE Cytogenet. Cell Genet. 87 (3-4), 191-194 (1999)
REFERENCE 20169181
2 (bases 1 to 2307)
Castagnola,P.
AUTHORS Direct Submission
TITLE Submitted (12-JUN-1998) Castagnola P., Biotechnologie in Oncologia,
JOURNAL Ist. Naz.le per la Ricerca sul Cancro, L.go R. Benzi 10 Genoa,

FEATURES	16132, ITALY
Source	Location/Qualifiers
	1..2307
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="Contig of pcphn2 and I.M.A.G.E. cloneID 1090695"
	/tissue_type="fetal brain"
	/dev_stage="19-23 weeks-old"
	12..89
sig.peptide	/gene="casp"
	12..1217
	/gene="casp"
	/codon_start=1
	/product="cartilage-associated protein (CASP)"
	/protein_id="CA07054.1"
	/db_xref="GI:3687322"
	/db_xref="SPRMBL:075718"
	/translation="MEPGRCRAALLALLVCALRAGRAQYERYSFRSPPEDEMLP ESAIRHALDTSGEHMAESVGYLETISRLNHLSDAFCHRNCSAPQEPAPGLAS YPELIRLFGFLIRAHCKLRCKQGI.PATRQSOPSREVIADQREPPYEFIDFAFKANN LPKIAAHTFLRLNHPDEMKRNMAYKSLPGAEYIKDLETKSYESLFIKAVRAN GEMNTSTIDMELALPDFFKAYPECLACGSRGRIKDFKYLSIADHYVEVECKIO CEENLTVPVIGGYPEVKFVATMYHYILOFAAYKILNDKNAACAVSYLLFDNDKMOON LVVYQYHRDTWGLSDHEHQPRPEAVQGFENVTTQCKELYDPAKENIMDDGEVEEYVD DLPLEFTS"
	12..1217
gene	/gene="casp"
	90..1214
mat.peptide	/gene="casp"
	/product="cartilage-associated protein (CASP)"
BASE COUNT	545 a 633 c 548 g 578 t 3 others
ORIGIN	

[illegible]

QY	539	tgaagaaatgataagagaaacatgtaataataaagacctgctgtgtccgaggacta	652
Db	542	tgaccaaatgatgaaagaaacaatgacgatattatgaagagcttccctggtgcccagagacta	601
QY	653	catlaaagaccttgaaaccaagtaataaagaaacctgttaatacccgagatgcggagata	712
Db	602	catttaaacaccttgaaacaaagctatattgaagacctgtttatccagacagtcgggacata	661
QY	713	caacggtgaagacctggaagaatccatacaagaacatgtagctgtgcccctccgactctt	772
Db	662	caacgctgagaacctggaagaaacatcatcacagacatgagactgtgcccctccgactctt	721
QY	773	caagccctttagagatgtctcgaagctgcgagggttccagggagataaagacttaaa	832
Db	722	caaacctcttttagcagatgtctccgagcttcgagggttccagggagatcaaacgacttcaa	781
QY	833	ggattctaaccttccataagaaataataaagtagtcttgaaatgcaaaatacaagt	892
Db	782	gsatttttacctttccatagagatcatattgtatagaattcttgaaatgcamaaatagactg	841
QY	893	tgaagaagaacctcaaccccaagttataggagctataccggttgagaattgtgttaccat	952
Db	842	tgaagaagaacctcaaccccaagttataggagctattccggttgagaattgtgttaccat	901
QY	953	gtataatcacttgtagctgttgcctataaagttgaagcgcgaagaatgtagccccctg	1012
Db	902	gtatcatattacttgtagctgttgcctattttagatttgaagcgcgttgaagatgtagccccctg	961
QY	1013	tgcagtcagctatactgtctcttcttgataagaaatgaagaagtcatactgtagcagaacctgtgta	1072
Db	962	tgcagtcagctatctgtctcttcttgcataagaaatgaagaagtcatactgtagcagaacctgtgta	1021
QY	1073	ttacacagatcaacaagaaggaacaccttggggccctctcgataagagaccttccaaagccaaagccga	1132
Db	1022	ttacacagatcaacaagaaggaacaccttggggccctctcgatagagaccttccaaagccaaagccga	1081
QY	1133	agcagttcagttcttcttaatgltgaaccaactccagaagaagctgtatgacttgtctaaga	1192
Db	1082	agcagttcagttcttctttaatgtgacacacacccagaagagcgtatgtgacttgtctaaga	1141
QY	1193	aaataataatgataatgataagaggagaagtgtggaataatgtgagtgacctcttggaact	1252
Db	1142	aaataataatgataatgataagaggagaagtgtggaataatgtgagtgacctcttggaact	1201
QY	1253	ggaggaagacccagctagcccaacgaacaaagaagaaactcctctgtgcgtctcagaagaaac	1312
Db	1202	ggaggaagacccagctagcccaacgaacaaagaagaaactcctctgtgcgtctcagaagaaac	1261
QY	1313	agattcttctgtcctttcccaacagcccagagctgtgtgataacctcaagacctctcttaac	1372
Db	1262	agattcttctgtcctttcccaacagcccagagctgtgtgataacctcaagacctctctttac	1321
QY	1373	tctccaagaatgaagaaggaaagcccgcgtctctcaactgaatgataatcaagaaggtagagctg	1432
Db	1322	tctccaagaatgaagaaggaaagcccgcgtctctcaactgaatgataatcaagaaggtagagctg	1381
QY	1433	ccttctcatcttccacacccctgcaacctcatglttcaacacctatcttctcaacctttttt	1492
Db	1382	ccttctcatcttccacacccctgcaacctcatglttcaacacctatcttctcaacctttttt	1441
QY	1493	gagatgtagtctgcgtctctctgtgccagagctcgagatgcaatgtaagcttctcaagctcaagt	1552
Db	1442	gagatgtagtctgcgtctctctgtgtgccagagctcgagatgcaatgtaagcttctcaagctcaagt	1501
QY	1553	caacctcgcctctgtggttcaagaacatctgtgtatcaagctcccgagtaacctgagat	1612
Db	1502	caacctcgcctctgtgtggggttcaagaacatctgtgtatcaagctcccgagtaacctgagat	1561
QY	1613	tacagcgaatgtgcaacacagcccggtcaatttgtatctta--gtagagacgggggtt	1670
Db	1562	tacagcgaatgtgcaacacaccccggtcaatttgtattttttatgaatgaagacgggggttt	1621

QY	1671	gccatcttggccaaggtctgctcgaaccttgaactcaagaatgcatcgtctt-9gac	1729
QY	1671	gcaatcttggccaaggtctgctcgaaccttgaactcaagaatgcatcgtctt-9gac	1729
Db	1622	gccatcttggccaaggtctgctcgaaccttgaactcaagaatgcatcgtctt-9gac	1681
QY	1730	tcccaaatgtcgcggagattcaagcgcttgagcaacatgcccggcctcttctcaaccttac	1789
Db	1682	tcccaaatgtcgcggagattcaagcgcttgagcaacatgcccggcctcttctcaaccttac	1741
QY	1790	acctgtctcttatacctcaacatcgtcttctcaacctcaacctcaacctcaacctca	1849
Db	1742	acctgtctcttatacctcaacatcgtcttctcaacctcaacctcaacctcaacctca	1801
QY	1850	caatgtctctcctcaatgcttcaagctcctcttcaacctcaacctcaacctcaacctca	1909
Db	1802	caatgtctctcctcaatgcttcaagctcctcttcaacctcaacctcaacctcaacctca	1861
QY	1910	tctctggtctgtcttcttcttcttcccaaaaaaacagtaattctttaaacaagaanaa	1969
Db	1862	tctctggtctgtcttcttcttcttcccaaaaaaacagtaattctttaaacaagaanaa	1921
QY	1970	catctcctagaagaatgataatgttgaaacatcctcttggtcttcttcttcaaatltt	2029
Db	1922	catctcctagaagaatgataatgttgaaacatcctcttggtcttcttcttcaaatltt	1981
QY	2030	atctcctcttctcctccatcccggaagaagatgtggaagaagatagcttaattctccagcc	2089
Db	1982	atctcctcttctcctccatcccggaagaagatgtggaagaagatagcttaattctccagcc	2041
QY	2090	tcaaatgtctctcaactgtgtctgactgttacaattctcagaccacacagaaacagat	2149
Db	2042	tcaaatgtctctcaactgtgtctgactgttacaattctcagaccacacagaaacagat	2101
QY	2150	tgaatgagagatgtagaatgacgaatgtgagcttcttcccaacttgcattcccaaat	2209
Db	2102	tgaatgagagatgtagaatgacgaatgtgagcttcttcccaacttgcattcccaaat	2161
QY	2210	acaagagtttggccgagctcccatcttgagagacaatgcttaagtaagctcccgagttgaa	2269
Db	2162	acaagagtttggccgagctcccatcttgagagacaatgcttaagtaagctcccgagttgaa	2221
QY	2270	aggaagaagacccagagctgtctgactgaattcaattcttcaacttcaagtaatttgaagta	2328
Db	2222	aggaagaagacccagagctgtctgactgaattcaattcttcaacttcaagtaatttgaagta	2280
RESULT 2			
AK001634			
LOCUS		1961 bp	linear PRI 22-FEB-2000
DEFINITION	Homo sapiens cDNA FLJ10772 fis, clone NT2RP4000243, highly similar		
ACCESSION	AK001634		
VERSION	AK001634.1	GI:7023007	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens testicular carcinoma cell line:NT2 cDNA to mRNA, clone_11b:NT2RP4 clone:NT2RP4000243.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (sites)		
TITLE	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Sasaki,N.		
JOURNAL	NEDO human cDNA sequencing project		
REFERENCE	2 (bases 1 to 1961)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-FEB-2000) to the DDBJ/EMBL/Genbank databases. Takao		
COMMENT	Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hrl.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)		
	NEDO human cDNA sequencing project supported by Ministry of		

[illegible]

QY 744 gacatgagcctggccctcccgactctctcaagcctttcaagtgctctgcagccgc 803
 DB 721 GACATGAGCCTGGCCCTCCGACTCTTCAAGCCCTTTTACGATGCTCTGCACCTGC 780
 QY 804 gaaggttccaggaatcaagactcaagattctacattccatccatgagatcatat 863
 DB 781 GAGGTTCCAGGAGATCAAGACTTCAAGATTCTTACCTTCCATAGCAGATCATAT 840
 QY 864 gtagaagltctggaatgcaaaataacagtgtagaagaaccccaacagttatagaagc 923
 DB 841 GTAGAAGTCTGGAATGCAAAATACAGTGAAGAAACCTCACCCAGTATATAGAGGC 900
 QY 924 taccgtttgagaatttgggtaccatgcatatcttaacttgcagttgacctattaa 983
 DB 901 TATCCGTTGGAATTTTGGCTACATGATCATTTCTTGACATTTGCTTATTAAG 960
 QY 984 ttgaagcacttgaagaagcagcccccgtgtagtcaagctatctgctctttgataagat 1043
 DB 961 TTGAAGCAGCTGAAGAATGCAAGCCCTGTGCAAGTCACTATCTCTTTGATCAGAA 1020
 QY 1044 gacaagltcatgcaagcagaacctgggtatattacagttacacagggagacattggggcctc 1103
 DB 1021 GACAAGTCAATGACAGCAAGACCTGTGATTAACAGTACCAAGAGGACAGTGGGCTC 1080
 QY 1104 tcgagatgcaacttccagcccgactgaagcagttcagttcttaagttagaccactc 1163
 DB 1081 TCAGATGAGCACTTCCACCCAGACTGAAGCAGTTCAAGTCTTAAATGTGACCACTC 1140
 QY 1164 cagaagagctgtatgacttctgaagaaataataatgatgatgagtagaggaagatt 1223
 DB 1141 CAGAAGAGCTGTATGACTTTTGCTAAGAAATAATATGATGATGATGAGGAGAAAGTT 1200
 QY 1224 gtggaatgtgtagtacctcttggaaacttgagagagaccagtagccacagcaacaa 1283
 DB 1201 GTGGAATGTGGAAGACCTTGTGAACGTGAGGAGACAGTACCCACAGCAACAAA 1260
 QY 1284 gagacttctctggcgttcaaggaacacagattctgtctcttcccaaacagccag 1343
 DB 1261 GAGACTTCTCTGGCGCTTCAGGAACACAGATTCTTGTCTTTCCCAACAGCCAG 1320
 QY 1344 ctgtgatacctcagagacctctctcttactcctcaagtgaaaggaaagcccgctctc 1403
 DB 1321 CTGTGATACCTCAGAGCCCTTCTCTTACTCTCCAAAGTGAAGGAAAGCCCTCTCT 1380
 QY 1404 ctaactgatatcatcaaggggtgagccgtctctctactcttaactgcaactctg 1463
 DB 1381 CTAACGTGATGATCAGAGGGGAGACCTGCTTCTCAATTCACACCTGCACCTCAATG 1440
 QY 1464 ttcaacacatcttctcaaccttttttttgaatgagtgctcgtctcttgcagcgctg 1523
 DB 1441 TTCACACTATCTTCTCACTTTTGTGAGATGAGATCTCGCTCTTGCCCAAGCTG 1500
 QY 1524 gagtgcgaatggcaagtlctcagctcactgcaacctcgccctcttgggttcaagaactct 1583
 DB 1501 GAGTGCATGACAGCTTCTCAGCTACATGCAACCTCCGCTTGGTGAAGCAATCT 1560
 QY 1584 gctgcataagcctcccgagatcctggagattacagggcatgtgcacacagcccgctaat 1643
 DB 1561 GCTGCATGACCTCCCGAGTACTGGGATACAGGCAATGTCACACAGCCCGGCTAAT 1620
 QY 1644 ttgatattttagtagaagcggggtttgcatgtagtggcagagcgtagtctgaactctga 1703
 DB 1621 TTGATATTTTATGATGAGAGGGGTTTTCAGATGTTGGCCAGGCTGGTCTCAACTCTTGA 1680
 QY 1704 cttaagatgcatccttctctggtccctcccaagtgctggagattacagcgctgaagccac 1763
 DB 1681 CTTCAAGATGATCATATCTCTTGGCTCTCCACAGTGTGGGATTAACAAGCGTGAACACC 1740
 QY 1764 atgcccggcctcttctcaacttcaacctgtctcttatacctcaacatctgtttcacac 1823
 DB 1741 ATGCCCGGCTCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1800
 QY 1824 cttaacccctgtcttcccaagttcaacactgtcttcccaagtttcaacagctgcttct 1883

DB 1801 CTTCACTCCCTGCTTCTCATGTTCACTGCTTCTTCCCATGTTCAATAGCTGCTTCT 1860
 QY 1884 taccatttgggttgaagggcagctctctgagctgtttttttttttttcccaagaaat 1943
 DB 1861 TACATTTTGGTGTGAAGGCGAGCTCTCTGCGCTGTGTTTTTTGTTTCCCAAGAAAT 1920
 QY 1944 cagttatattttaaataagaanaaacattcctgaagatg 1984
 DB 1921 CAGTATTTATTTTAAATTAAGAAAAATCTCTTCAAGATG 1961
 RESULT 3
 LOCUS BC008745
 DEFINITION Homo sapiens, cartilage associated protein, clone MGC:1926
 IMAGE:3347384, mRNA, complete cds.
 ACCESSION BC008745
 VERSION BC008745.1 GI:14250580
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1439)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland,
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nsl.nih.gov
 Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-T., Karlins,E., Legaspi,R., Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stantiripop,S., Thomas,P.J., Tjongson,E.E., Touchman,J.W., Tsurugon,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: 1 Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3687321.
 FEATURES
 source
 location/Qualifiers
 1..1439
 /organism="Homo sapiens"
 /db_xref="locusid:10491"
 /db_xref="taxon:9606"
 /clone="MGC:1926 IMAGE:3347384"
 /tissue_type="Kidney, renal cell adenocarcinoma"
 /clone_1lb="NIH_MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 36..1241
 /codon_start=1
 /product="cartilage associated protein"
 /protein_id="AAH08745.1"
 /db_xref="gi:14250581"
 /translation="MEPGRGAALLLALCVACALRAGRAOYERYSPRSPDELMLP
 EASVRLADKYSGEHMAESGYEISRLRLHLDSEAFCRNCSAPOPAPAGLAS
 YPELRALFGILRRHNLCKRCKGIPARROSPSDVADPQRRPYFLQFAFYRKAN

PROTEIN	GENE	BASE COUNT	ORIGIN
<p>/protein_id="CAA07053.1"</p> <p>/db_xref="GI:3687320"</p> <p>/db_xref="SP:SPRMBL:088698"</p> <p>/translation="MGPRSPAAALVLLVCGCAATPGRCQGEYRSPFNPRDEIMPLE SAYNAHLDQYSGEHMAISVCTLEYSILHLHLRLHRLSEAFCHRNCSAATPAPACAPASH AELRLGCVTLRRACQLKRCQGLPAFQKQSQSVLADPQOREPRTKLOAFYANL PKALAAAHYTLHKHPDDEMMKRNMEYRKSLEPGADHDKLETKSESLFYRAVAYANG ENMTTISIDMELAPFLKAFYECLACESRLEIKDKRDEYLSIADHYVLEQKINC EEWLTPIVIGGYPVEKEFATMYHYLOFAFYKYKLINDLKNAKPAVSYLEFDQSDRWQML VVYYVHRDKMGLSDSEHFQRPPEAVQFNVTTLOKELYDFAQEHLMDDDESEVEYVDD LLRETESA"</p>	<p>gene</p> <p>mat_peptide</p> <p>72..1274</p> <p>/gene="casp"</p> <p>147..1271</p> <p>/gene="casp"</p> <p>/product="cartilage-associated protein (CASP)"</p>	<p>362 a</p> <p>503 c</p> <p>449 g</p> <p>371 t</p>	

Query Match	24.4%	Score 944	DB 10	Length 1685
Best Local Similarity	83.6%	Pred. No. 1.2e-150		
Matches 1080; Conservative	2;	Mismatches 207;	Indels 3;	Gaps 1;

[illegible]

Db	728	GAACAGGAAAGCTCCATTTCGACATGGAGGCTCGGCTTCCGACTTCCTCAAGACCTT	787
QY	782	ttacagatgctcgcagccctgcgaaggtctccaggagataccaaggacttaaggattcta	841
Db	788	CTACAGAGTCCCTGGGCTGCTTGAGGGGCGGGGGAGTCAAGAGACTTCACAGACTTCA	847
QY	842	ccttccatagagagatcatctatgtagaattctcggaaatgcacaataacagtgtgaaagaga	901
Db	848	CCTGTCCATAGCAGATCACTATGTGGAAGTTTGTGAGTAAAGATTCTGTGTGAGAGAC	907
QY	902	cctcacccccagttatagagagcctacccggttgagaattgtggtctaccatglatcata	961
Db	908	CCTCACCCCAAGTCATAGAGGGCTATCCCTGGAGAAATTTGTGGGCACCATGTACACATCA	967
QY	962	cttgaagtttgctcttataagtttaagcacttgaagatggagacccccctgtgcagttcg	1021
Db	968	TTTACAGTTTGGCTTTATACAAAGTTGAATGATGTGAAGATGAGACCCCTGTGCGCTCAG	1027
QY	1022	ctatctgctctcttgcagaaatgacaaggctcatgcaacgaaccttgglytlatlaccagta	1081
Db	1028	CTACCTGCTCTTTGACACAGAGTGACAGGGGTATGCAACAGAACCTGTGTACTATACATCA	1087
QY	1082	ccacaggagacacttgyggccctctcgatgtagaaccttccagccccagacctgaaagcagttca	1141
Db	1088	CCACCGGGGCAAGTGGGGGCTCTCGAGTAGACACTTCCAGCCCAAGCCCGAAGCAAGTTCA	1147
QY	1142	gtctcttaatgtgacacacactccagaagagcttgtatgactcttgcataagaaataataat	1201
Db	1148	GTTCTTTATATGTGACGAGCCCTCCAGAAAGAACTGTACAGACTTCGCTCAAGAAACACTAT	1207
QY	1202	ggaatgatgtagaggagaagattgtggaatataatgtgataatgacactctcttggaaatggaagagac	1261
Db	1208	GGATGACGATGAGGAGAGAGTGTGTGAGATGTGAGACGACTTGTGGAGACACGGAAGATGC	1267
QY	1262	cagctagccacacagcaaccaaagagaactccct	1293
Db	1268	TGCTTAGTCCACAGAGGCTTAAGGAACCTTCT	1299

LOCUS	GGCARRP	1665 bp	mRNA	linear	VRT 26-JUN-1997
DEFINITION	GGCARRP				
ACCESSION	X.9allus mRNA for cartilage associated protein.				
VERSION	X97607				
KEYWORDS	X97607.1 GI:1296525				
SOURCE	cartilage associated protein.				
ORGANISM	chicken.				
REFERENCE	Gallus gallus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus				
TITLE	1 (bases 1 to 1665)				
JOURNAL	Castagnola, P., Gancetti, M., Morello, R., Tonachini, L., Martin, O., Gaggero, A. and Candada, R.				
MEDLINE	Cartilage associated protein (CASP) is a novel developmentally regulated chick embryo protein				
REFERENCE	J Cell. Sci. 110 (Pt 12), 1351-1359 (1997)				
AUTHORS	97360293				
TITLE	2 (bases 1 to 1665)				
JOURNAL	Castagnola, P.				
MEDLINE	Direct Submission				
REFERENCE	Submitted (22-APR-1996) P. Castagnola, Centro di Biotecnologie Avanzate, Differenziamento Cellulare, L.go R. Benzi, 10 Genova, 16132, ITALY				

FEATURES	
source	Location/Qualifiers
	1..1665
	/organism="gallus gallus"
	/db_xref="taxon:9031"
	/cell_type="Chondrocyte"
	/dev_stage="embryo"
	/lab_host="E.coli"
CDS	43..858

CDS

43.858

[illegible]

QUERY MATCH	SCORE	DB	LENGTH
12.1%	467	DB 9	1889
Best Local Similarity	64.1%	Pred. No. 9.7e-70	

[illegible]

QY	1132	aagcgcattcagtlctttaa	tgtgaccaccc	caggagaagctgtatgactt	gtctaagg	1191
Dd	1110	AGCCATGCTCTTACACA	CAACGACCGCCGACCTGGGGAGCCTGCCTGGAGTTTCACCACAA	1169		
QY	1192	aaataataatgatgatgatga	ggaggaagctgttgga	1228		
Dd	1170	TGTACTGCAGTCAGTGATG	GATGAGATGAGAGCTGGAGA	1206		
RESULT	7					
LOCUS	BC007942					
DEFINITION	BC007942	2037 bp	mRNA	linear	-PRI 12-JUL-2001	
ACCESSION	BC007942	Homo sapiens, nucleolar autoantigen (55kD) similar to rat synaptonemal complex protein, clone MGC:14267 IMAGE:4130726, mRNA, complete cds.				
VERSION	BC007942					
KEYWORDS	MGC.	GI:14044027				
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
TITLE	1 (bases 1 to 2037)					
JOURNAL	Strausberg,R. Direct Submission Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
COMMENT	Contact: MGC help desk Email: cgabbs@email.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Shavchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R., Lin,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C., McConwell,J.J., Pearson,R., Snyder,B., Stantiripop,S., Thomas,P.J., Tongelson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.					
FEATURES	Source					
CDS						

[illegible]

RESULT	11
LOCUS	HSU47621
DEFINITION	2347 bp mRNA linear PRI 23-OCT-1997
ACCESSION	U47621 Homo sapiens nucleolar autoantigen Noss mRNA, complete cds.
VERSION	U47621.1 GI:1491808
KEYWORDS	
SOURCE	human.

[illegible]

